



XSLMFCE

16-19
Nov
2021

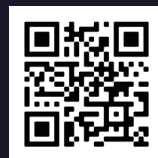
Actas del X Congreso de la Sociedad de Lógica, Metodología y Filosofía de la Ciencia en España

Hospedería Fonseca, Universidad de Salamanca

Conferencia LULLIUS

Sandra Mitchell

(University of Pittsburgh)



VNIVERSIDAD
D SALAMANCA

CAMPUS DE EXCELENCIA INTERNACIONAL

INSTITUTO
eCvT
INSTITUTO UNIVERSITARIO DE ESTUDIOS
DE LA CIENCIA Y LA TECNOLOGÍA



Actas del X Congreso de la Sociedad de
Lógica, Metodología y Filosofía de la
Ciencia en España

Salamanca, 16-19 de noviembre de 2021

Proceedings of the X Conference of the
Spanish Society of Logic, Methodology
and Philosophy of Science

Salamanca, 16-19 November 2021

EDITADO POR:

Ana Cuevas (USAL)
Obdulia Torres (USAL)
V́ctor Aranda (USAL)
Andrei Moldovan (USAL)

MAQUETACIÓN: Intergraf

DISEÑO DE LA PORTADA: Rodrigo López Orellana

ISBN: 978-84-09-35510-5

COMITÉ ORGANIZADOR USAL

Obdulia M. Torres
Ana Cuevas Badallo
Andrei Moldovan
Esther Palacios Mateos
Rodrigo López Orellana
Daniel Labrador Montero
Álvaro Rodríguez
Benedicto Acosta
Víctor Aranda
José Javier González López
Mariano Martín Villuendas
Ángel Álvarez Marcos

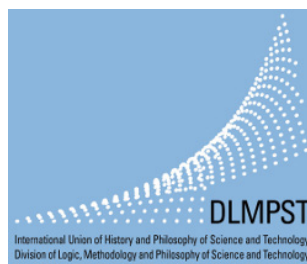
COMITÉ ORGANIZADOR SLMFCE

Cristina Corredor
José Martínez Fernández
Paula Olmos
Fernando Soler
Arantza Etxeberria
María José García Encinas
Sergi Oms
David Pérez Chico

COMITÉ CIENTÍFICO

M.J. García Encinas
María Caamaño
Marta Campdelacreu
Sergi Oms
Fernando Soler
Jose Díez
Arantza Etxeberria
Andrei Moldovan
Adán Sus
Obdulia Torres
Ana Cuevas
David Pérez Chico
Fernando Soler
Mathieu Fontaine
Henrik Zinkernagel
Paula Olmos
Ángeles Jiménez Perona
Cristina Corredor
Iñaki San Pedro
José Martínez
Javier Vilanova Arias
Jesús Vega

INSTITUCIONES PATROCINADORAS



Índice general

| | |
|--|-----------|
| Prefacio | XVIII |
| Premio para Jóvenes investigadores e investigadoras | XX |
| 2021 Lullius Lectures <i>Sandra D. Mitchell</i> | XXI |
| Argumentación | 1 |
| 1 Burrieza, Alfredo y Yuste-Ginel, Antonio <i>Formal argumentation and epistemic logic: what can they do for each other?</i> | 2 |
| 2 Casanova, Assumpció <i>Obtención de explicaciones orientadas al humano a partir de demostraciones automáticas de teoremas: un algoritmo de transformación</i> | 6 |
| 3 Haro Marchal, Amalia <i>The interactionist approach to speech acts and its role in the analysis of argumentation</i> | 10 |
| 4 Lewinski, Marcin <i>Challenging authority with argumentation: The pragmatics of not accommodating authority</i> | 14 |
| 5 Migura, Fernando <i>Un argumento falaz para defender el multiverso frente al diseño inteligente</i> | 17 |
| 6 Ruiz Rincón, Diana Lizbeth y González Rivas, Francisco Javier <i>De la argumentación en espacios emergentes, una propuesta entre el juego y la práctica semántica</i> | 21 |
| Lógica, Historia y Filosofía de la Lógica | 25 |
| 7 Aranda, Víctor <i>El primer Tarski y la concepción universalista de la lógica</i> | 26 |
| 8 Barés Gómez, Cristina <i>An Analysis of Ancient Medical-Veterinary Diagnosis</i> | 30 |

| | | |
|----|--|----|
| 9 | Barés Gómez, Cristina; Nepomuceno-Fernández, Angel y Salguero-Lamillar, Francisco J. | 33 |
| | <i>Methodological Analysis for the Reconstruction of Language Families. The case of Semitic Languages</i> | |
| 10 | Belastegui, Javier | 36 |
| | <i>Essentialist Negative Kinds and Internal Negations</i> | |
| 11 | Benito, Carlos y Martínez, José | 41 |
| | <i>A solution to the problem of mixed inferences by juxtaposition</i> | |
| 12 | Canela Morales, Luis Alberto | 44 |
| | <i>La construcción del espacio lógico de posibilidad</i> | |
| 13 | Carrete Varela, Pablo | 49 |
| | <i>Igualitarismo y Realismo en la metaontología de Hirsch</i> | |
| 14 | Conti, Ludovica | 53 |
| | <i>Abstraction's Logicality and Invariance</i> | |
| 15 | Fontaine, Matthieu | 57 |
| | <i>Cross-Fictional Identity in the Artifactual Theory of Fiction</i> | |
| 16 | García-Arnaldos, María Dolores | 59 |
| | <i>Diálogo en metafísica</i> | |
| 17 | González López, José Javier | 63 |
| | <i>Una traducción de la teoría de tipos proposicional (PTT) a la lógica multivariada de primer orden (MSL)</i> | |
| 18 | Gracia Di Rienzo, Alejandro | 65 |
| | <i>Una vindicación del conjunto universal</i> | |
| 19 | Verdeé, Peter; Sain-Germier, Pierre; Dantas de Oliveira, Joao Daniel y Terrés Villalonga, Pilar | 71 |
| | <i>Constitutional logics as hyperintensional representations of extensional logics</i> | |
| 20 | Zardini, Elia | 76 |
| | <i>The Final Cut</i> | |

| | |
|---|-----|
| Filosofía de la Mente y Epistemología | 77 |
| 21 Barbarrusa, Daniel y Medina Vizueté, Lola <i>Teorías de la conspiración: explotación, producción y tergiversación de injusticias epistémicas</i> | 78 |
| 22 Brugnami, Enrico <i>Alcances y límites de la noción de punto de vista desde la teoría de los espacios conceptuales</i> | 83 |
| 23 Camarena Bononad, Joan <i>El neo-cartesianismo como callejón sin salida: cómo la psicología ecológica se acomoda mejor en las ciencias cognitivas</i> | 86 |
| 24 Cañas Muñoz, Manuel Arnulfo <i>Contraejemplos a la solución fiabilista del problema de la inducción</i> | 90 |
| 25 Fairhurst, Jordi <i>Rethinking the epistemology of deep disagreements through the lens of Wittgenstein's later philosophy</i> | 95 |
| 26 Fernández, Víctor y Núñez de Prado Gordillo, Miguel <i>Mental Disorders, Values and Expressivism</i> | 100 |
| 27 Galadí García, Javier; Langa Rosado, José Antonio y Soler-Toscano, Fernando <i>Información y conciencia</i> | 103 |
| 28 Gómez-Ledo, Ignacio <i>Nuevos desarrollos en torno a la posibilidad de una Epistemología feminista de los goznes</i> | 108 |
| 29 González de Prado Salas, Javier <i>Good players in epistemic games</i> | 112 |
| 30 Hermoso Durán, Juan <i>Reducción y eliminación. Una respuesta a Strawson (2019)</i> | 116 |
| 31 Hernández-Conde, José V. <i>¿Por qué un concepto aprendido no puede ser invariante?</i> | 118 |
| 32 Ibarrodo Murguialday, Julen <i>Racionalidad y excusas</i> | 123 |

| | | |
|----|---|------------|
| 33 | Jaume, Andrés Luis | 127 |
| | <i>Escepticismo, relativismo y puntos de vista: una defensa del perspectivismo alético y meliorista</i> | |
| 34 | Liz, Manuel | 131 |
| | <i>Verdad relativista y verdad perspectivista</i> | |
| 35 | Monsó, Susana | 136 |
| | <i>Playing possum: Animal thanatosis and the concept of death</i> | |
| 36 | Navarro, Jesús | 140 |
| | <i>Secrets: An Epistemological Account</i> | |
| 37 | Núñez de Prado Gordillo, Miguel y Almagro Holgado, Manuel | 144 |
| | <i>Yet another defence of doxasticism about delusions</i> | |
| 38 | Pérez Otero, Manuel | 148 |
| | <i>Etiología doxástica y proposiciones lotería</i> | |
| 39 | Pino, Dani | 152 |
| | <i>Collective Understanding-Why</i> | |
| 40 | Serrahima, Carlota y Martínez, Manolo | 156 |
| | <i>Painful Moods and the Affective Component of Primary Dysmenorrhea</i> | |
| 41 | Vera Vega, Pablo | 160 |
| | <i>Epistemología del desinterés por la verdad</i> | |
| | Filosofía del Lenguaje | 163 |
| 42 | Almagro, Manuel | 164 |
| | <i>Too much confidence in certain beliefs. Affective polarization and the possibility of error in ascribing certain mental states</i> | |
| 43 | Baumgartner, Lucien | 168 |
| | <i>The Dual Character Concept Approach to Normative Generics</i> | |
| 44 | Delgado, Laura | 170 |
| | <i>The Logic of Polyreference</i> | |
| 45 | Domínguez Armas, Álvaro y Soria Ruiz, Andrés | 172 |
| | <i>Provocative insinuations</i> | |

| | | |
|----|---|-----|
| 46 | Fernández, Ainhoa | 177 |
| | <i>The Classificatory Model of Communication and Indexical Utterances</i> | |
| 47 | Geirsson, Heimir | 180 |
| | <i>Singular Thought and Strong Acquaintance</i> | |
| 48 | Llorente, Carlota G. | 182 |
| | <i>El problema de la dimensión ontológica en la hipótesis de la relatividad lingüística</i> | |
| 49 | Macià, Josep | 186 |
| | <i>Grice and Davis on Meaning</i> | |
| 50 | Moreno Zurita, Alba | 189 |
| | <i>The effect of slurs: World-ordering and the impossibility of retraction</i> | |
| 51 | Pérez-Navarro, Eduardo | 192 |
| | <i>The way things go: Moral relativism and suspension of judgment</i> | |
| 52 | Picazo Jaque, Claudia; Crespo, Inés y Heise, Andreas | 196 |
| | <i>Metaphor identification via contextual coherence</i> | |
| 53 | Reuter, Kevin; Baumgartner, Lucien y Willemsen, Pascale | 200 |
| | <i>Tracing Thick Concepts Through Corpora</i> | |
| 54 | Rychter, Pablo | 206 |
| | <i>Logical conventionalism, truthmaking and expressivism</i> | |
| 55 | Sambrotta, Mirco | 210 |
| | <i>Are empty names all the same?</i> | |
| 56 | Stojanovic, Isidora | 214 |
| | <i>Slurring Terms in Narrative</i> | |
| 57 | Viejo, José Manuel | 219 |
| | <i>Do direct reference theorists have an attitude problem?</i> | |
| | Historia de la Ciencia | 222 |
| 58 | Barroso, Milagros | 223 |
| | <i>Hacia una Teoría Consciente de la Representación Científica</i> | |

| | | |
|----|---|------------|
| 59 | Filomeno, Aldo | 226 |
| | <i>Emergent Laws in cosmology</i> | |
| 60 | Krauss, Alexander | 229 |
| | <i>How early humans made science possible</i> | |
| 61 | Labrador Montero, Daniel | 232 |
| | <i>Transferencia conceptual y el rol de la metáfora en ciencia: el caso de la noción de 'división del trabajo' en la biología del siglo XIX</i> | |
| 62 | Pascual Alba, Joseba | 236 |
| | <i>La (re)emergencia de las filosofías pluralistas de la ciencia: variaciones divergentes sobre un tema</i> | |
| 63 | Pérez del Amo, M. ^a de Los Ángeles | 241 |
| | <i>Sobre el quiebre del aparato conceptual kuhniano en el terreno de la biología evolutiva contemporánea</i> | |
| 64 | Vecchi, Davide | 244 |
| | <i>The philosophical significance of the molecular interpretation of evolution</i> | |
| | Filosofía y Metodología de la Ciencia | 248 |
| 65 | Airoidi, Giorgio | 249 |
| | <i>The arrival of the Smartest</i> | |
| 66 | Alonso, Ana María | 253 |
| | <i>Papel de los Valores en las Ciencias de Diseño: de la Axiología de la Investigación a la Ética de la Ciencia en Educación</i> | |
| 67 | Amilburu Martínez, Alba y Soler Parra, Jaime | 257 |
| | <i>Vida como género natural: dificultades metafísicas en la definición de vida</i> | |
| 68 | Anta, Javier | 261 |
| | <i>Framing the Epistemic Schism of Statistical Mechanics</i> | |
| 69 | Arca, Carlos | 266 |
| | <i>¿Es la convexidad una condición necesaria de las propiedades (naturales) cromáticas en la ceguera al color? Un estudio de caso desde la filosofía experimental</i> | |

| | | |
|----|---|-----|
| 70 | Baravalle, Lorenzo | 270 |
| | <i>An agenda for the philosophy of biological computation</i> | |
| 71 | Beisbart, Claus y Reuter, Kevin | 273 |
| | <i>What is the folk concept of life?</i> | |
| 72 | Bich, Leonardo | 278 |
| | <i>Grounding minimal cognition in regulatory control</i> | |
| 73 | Corral-Villate, Amaia | 282 |
| | <i>On particle disappearance in classical mechanics</i> | |
| 74 | Cortés García, David y Etxeberria Agiriano, Arantza | 283 |
| | <i>Individualidad biológica y viviparidad: una investigación filosófica desde las transiciones evolutivas y la evo-devo</i> | |
| 75 | Denia, Elena | 288 |
| | <i>El experimento científico por simulación y el problema de la materialidad</i> | |
| 76 | García Portela, Laura | 291 |
| | <i>Attribution methods in climate science: exploring the relevant differences</i> | |
| 77 | García-Lapeña, Alfonso | 294 |
| | <i>Estimating closeness to the truth for scientific laws</i> | |
| 78 | Garrido Rodríguez, Natividad | 299 |
| | <i>La narración del paciente como “evidencia epistemológica” en medicina</i> | |
| 79 | Gefaell, Juan y Saborido, Cristian | 302 |
| | <i>The Extended Evolutionary Synthesis and Scientific Incommensurability. Taking Kuhn Seriously</i> | |
| 80 | González Herrerías, César | 307 |
| | <i>Caracterización del concepto de dato en los ácidos nucleicos desde un enfoque diafórico</i> | |
| 81 | González Martín, Joshua Alexander | 311 |
| | <i>Análisis de la condición epistémica de la responsabilidad moral desde “tres enfoques” para el contexto tecnológico</i> | |

| | | |
|----|--|-----|
| 82 | Jha, Aditya | 315 |
| | <i>Not So Distinctively Mathematical Explanations: Microphysics matters</i> | |
| 83 | Kvasz, Ladislav | 319 |
| | <i>Descartes on Mathesis Universalis and on Method</i> | |
| 84 | Lanza, Henar | 322 |
| | <i>Las metáforas científicas y nuestra concepción de la Tierra</i> | |
| 85 | López Mas, Roberto | 326 |
| | <i>Conocimiento científico y toma de decisiones: (in)certidumbre de la evidencia en la ciencia reguladora</i> | |
| 86 | Luján, José Luis y Todt, Oliver | 329 |
| | <i>Políticas epistémicas y controversias en ciencia reguladora</i> | |
| 87 | Luque, Victor J. y Baravalle, Lorenzo | 333 |
| | <i>A Pricean Foundation for Cultural Evolutionary Theory</i> | |
| 88 | Martín Villuendas, Mariano | 338 |
| | <i>¿Es posible ser pragmatista en ciencia? Una caracterización deflacionaria de la comprensión científica</i> | |
| 89 | Martínez, David Álvaro | 342 |
| | <i>Programas de Lakatos e investigación antiviral contra el SARS-CoV-2</i> | |
| 90 | Menatti, Laura y Saborido Alejandro, Cristian | 348 |
| | <i>Health: a new conceptualization for the post-pandemic scenario. When philosophy of medicine meets environmental philosophy</i> | |
| 91 | Monterde Fuertes, Alberto | 352 |
| | <i>Reductionism and holism in pain medicine: dehumanization of medicine and complexity as the justification of clinical ethics</i> | |
| 92 | Morato, Vittorio y Carrara, Massimiliano | 357 |
| | <i>Copies and dependence</i> | |
| 93 | Osorio, Javier | 361 |
| | <i>Disagreement and polarization in scientific communities</i> | |
| 94 | Pérez Escobar, José Antonio | 365 |
| | <i>There are no “genuine mathematical explanations” in biology, but mathematical heuristics</i> | |

| | | |
|-----|---|-----|
| 95 | Pérez-González, Saúl | 369 |
| | <i>Mechanisms and policymaking: lessons from EBM+</i> | |
| 96 | Perry, Eoin | 373 |
| | <i>Philosophical Foundations for an Empirical Bayesianism</i> | |
| | <i>Interpretation of Null Hypothesis Significance Testing in Psychology</i> | |
| 97 | Rivadulla, Andrés | 379 |
| | <i>Theoretical innovation and scientific explanation. Two intertwined tasks</i> | |
| 98 | Rodríguez, Álvaro | 385 |
| | <i>Minas a cielo abierto en la comarca de Olivenza: las controversias científico-tecnológicas públicas</i> | |
| 99 | Sánchez Gómez, Pedro J. y Suárez, Mauricio | 389 |
| | <i>An inferentialist account of Lewis structures</i> | |
| 100 | Sanjuán, Mariano | 394 |
| | <i>Pluralism beyond complexity: Reconceptualizing the disunity of Psychiatry</i> | |
| 101 | Solé, Albert | 398 |
| | <i>Contra la redundancia</i> | |
| 102 | Suárez, Mauricio | 402 |
| | <i>Evolutionary Fitness: Propensities, Probabilities, and Experimental Statistics</i> | |
| 103 | Suárez, Javier y Stencel, Adrian | 406 |
| | <i>Do somatic cells really sacrifice themselves? Altruism and coercion in the evolution of multicellularity</i> | |
| 104 | Vázquez, Margarita y Liz, Manuel | 409 |
| | <i>Perspectivismo científico</i> | |
| 105 | Villanueva, Enrique | 414 |
| | <i>La deuda técnica en el desarrollo de la inteligencia artificial y sus consecuencias</i> | |
| 106 | Villegas, Cristina | 417 |
| | <i>Modeling Evolutionary Propensities</i> | |

| | |
|--|------------|
| Ciencia, Tecnología y Sociedad | 421 |
| 107 González Rodríguez, María Dolores <i>El arma del telescopio</i> | 422 |
| 108 Muñoz, José Manuel <i>Dificultades en torno a la inclusión del libre albedrío como neuroderecho humano</i> | 426 |
| 109 Pérez Sedeño, Eulalia <i>Valores contextuales para una ciencia mejor</i> | 431 |
| 110 Requejo, Libia S. y González-Bravo, M. Isabel <i>Percepción e implicación social en la regulación de la ciencia y la tecnología</i> | 435 |
| 111 Rodríguez Navia, Alizon <i>En la búsqueda de la visibilidad y el reconocimiento: la producción científica de las mujeres en el Perú</i> | 441 |
| 112 Urueña López, Sergio <i>Anticipación e innovación responsable: los roles de las representaciones del futuro en la gobernanza científico-tecnológica</i> | 445 |
| | |
| Symposium: Ontological and Metaontological Deflationisms | 450 |
| 113 De Donato-Rodriguez, Xavier y Falguera, José L. <i>Carnap's metaontological deflationism and some neo-Carnapian approaches</i> | 453 |
| 114 Azzano, Lorenzo <i>Truth is not enough. How to be a semantic and existence deflationist, without being a metaontological deflationist</i> | 456 |
| 115 Martínez Vidal, Concha <i>Rayo's view (2013) from a Deflationist perspective</i> | 459 |
| 116 García-Carpintero, Manuel <i>Fictionalism about Fictional Characters (Reloaded)</i> | 462 |
| 117 Ponte Azcárate, María <i>Indexicals, the self and time</i> | 465 |

| | | |
|--|--|------------|
| Symposium: Philosophies of Scientific Experimentation: Old and New Challenges | | 468 |
| 118 | Vallejos, Gabriel <i>Purificación de entidades y obtención de factores causales: aspectos epistemológicos de la práctica experimental en bioquímica</i> | 470 |
| 119 | López Orellana, Rodrigo <i>Modeling, explanation, and experimentation in evolutionary biology: the distinction between phenomena and target system</i> | 473 |
| 120 | Jiménez-Buedo, María <i>Reactividad y Validez experimental</i> | 475 |
| 121 | Esposito, Maurizio <i>La filosofía del experimento y la teoría poiética del conocimiento: ¿Una relación constitutiva?</i> | 477 |
| Simposio: Cuestiones de Metafísica Hoy | | 480 |
| 122 | García-Encinas, María José <i>Sobre la no-existencia de las “entidades” de ficción</i> | 483 |
| 123 | Campdelacreu, Marta <i>Sobre la solución de Fine al problema de la fundamentación</i> | 486 |
| 124 | González, Joan <i>Superveniencia de propiedades económicas de los objetos a partir de propiedades físicas: análisis y limitaciones</i> | 489 |
| 125 | Echavarría, Agustín <i>Metafísica modal contemporánea y metafísica aristotélica: ¿continuidad o ruptura?</i> | 492 |
| 126 | Franco, Reynner <i>Identidad errática y uso de la intuición</i> | 495 |

| | |
|---|----------------|
| Simposio: Pragmática de las Razones | 496 |
| 127 Alhambra Delgado, José | 498 |
| <i>Argumentación por analogía y paridad de razones</i> | |
| 128 Marraud, Hubert | 501 |
| <i>La construcción dialógica de las razones</i> | |
| 129 Olmos, Paula | 504 |
| <i>¿Qué significa “reconocer razones”?</i> | |
| 130 Urbina, Raúl | 507 |
| <i>Pragmática de la publicidad. El “amor” por las marcas y la creación de una comunidad discursiva como estrategia argumentativa</i> | |
| Simposio: Prácticas Argumentativas | 510 |
| 131 Alonso, Enrique y Hernández, Soledad | 512 |
| <i>Diagramación argumentativa con herramientas visuales</i> | |
| 132 Corredor, Cristina | 516 |
| <i>Efectos perlocutivos de la argumentación</i> | |
| 133 Vega Reñón, Luis | 519 |
| <i>La singular práctica de la argumentación vía exemplum</i> | |
| 134 Vilanova, Javier | 522 |
| <i>Filosofía de la Lógica Ordinaria: la Lógica Informal como hija de la Filosofía del Lenguaje Ordinario</i> | |
| Simposio: Cuerpos Sexuados: Perspectivas Filosóficas sobre Teorías Biológicas de la Diferencia Sexual y la Identidad de Género | 525 |
| 135 Nuño de la Rosa, Laura y Triviño, Vanessa | 528 |
| <i>El sistema Sexo/Género como clase emergente: sobre la necesidad de una metafísica feminista</i> | |
| 136 Arraiza Zabalegui, Maite | 531 |
| <i>Múltiples devenires trans*: Un análisis transfeminista de las principales aportaciones de la neurobiología a la cuestión de las identidades trans*</i> | |

| | | |
|--|--|---------|
| 137 | Rodríguez, Ainhoa y Etxeberria, Arantza | 534 |
| | <i>Por qué quiero sangrar (y por qué tú también deberías): Función, normatividad y salud menstrual</i> | |
| 138 | Basanta, Silvia y Nuño de la Rosa, Laura | 537 |
| | <i>El orgasmo femenino y el concepto de homología</i> | |
| Simposio: Ontologías Híbridas: Objetos Científicos y Culturas Visuales entre la Industria, la Clínica y el Laboratorio | | 539 |
| 139 | Romero de Pablos, Ana | 541 |
| | <i>Significados y materialidades de un objeto híbrido: el uranio español en la década de los sesenta</i> | |
| 140 | Rego Robles, Miguel Ángel | 542 |
| | <i>Imágenes digitales del cerebro en la década de 1970 y la construcción de una tecnología híbrida</i> | |
| 141 | Sclavo, Daniela | 543 |
| | <i>Cocinas, bancos de semillas y laboratorios: los chiles como objetos científicos</i> | |
| 142 | Barahona, Ana | 544 |
| | <i>Redes colaborativas y objetos híbridos: mujeres genetistas en México de 1960 a 1970</i> | |
| 143 | Santesmases, María Jesús | 545 |
| | <i>El feto humano como ontología híbrida: cromosomas, ecografías y fotografías en la década de 1960</i> | |
| Simposio: Lenguaje y Humor | | 546 |
| 144 | Vilanova, Javier | 549 |
| | <i>El valor epistemológico del humor verbal: la risa como logro cognitivo</i> | |
| 145 | Blanco Salgueiro, Antonio | 553 |
| | <i>Relativamente gracioso</i> | |

| | | |
|-----|---|------------|
| 146 | De Salvador Agra, Saleta | 555 |
| | <i>La contestación de las dicotomías: la sonrisa de Austin</i> | |
| 147 | Duarte Calvo, Antonio | 558 |
| | <i>Siempre nos quedará la risa</i> | |
| | Symposium: The Context-Sensitivity of Expressives | 561 |
| 148 | Cepollaro, Bianca; Domaneschi, Filippo y Stojanovic, Isidora | 563 |
| | <i>Literally “a jerk”. An experimental study of the semantics of negative expressives</i> | |
| 149 | Carranza, Víctor | 567 |
| | <i>The pragmatics of all-purpose pejoratives</i> | |
| 150 | Berskyte, Justina | 570 |
| | <i>Jeremy is a... Expressive-Relativism and Expressive Predicates in Copular Positions</i> | |
| 151 | Baumgarten, Lucien; Reuter, Kevin y Willemsen, Pascale | 572 |
| | <i>Thick Ethical Concepts – The Extent and Possible Explanations of the Polarity Effect</i> | |

Prefacio

Ana Cuevas y Obdulia Torres

El X Congreso de la Sociedad de Metodología, Lógica y Filosofía de la Ciencia tuvo lugar en la Universidad de Salamanca entre los días 16 y 19 de noviembre de 2021 y estuvo organizado por el Instituto Universitario de Estudios de la Ciencia y la Tecnología y por el Departamento de Filosofía, Lógica y Estética de la Universidad de Salamanca.

La Sociedad de Lógica, Metodología y Filosofía de la Ciencia en España se fundó en el marco de su Primer Congreso que tuvo lugar en Madrid en diciembre de 1993. Desde un primer momento se planteó que la finalidad de la Sociedad sería fomentar la investigación e impulsar las relaciones profesionales de los miembros de las áreas de lógica, filosofía de la ciencia, e historia de la ciencia, así como de áreas afines en España. A partir de entonces los Congresos han cumplido una parte importante de esa finalidad, llevándose a cabo de manera ininterrumpida cada tres años.

En esta ocasión, después de haber pasado los peores momentos de la situación de pandemia por la Covid-19, nos pudimos reunir en Salamanca durante esos días y compartir ideas y discusiones como hemos venido haciendo desde la fundación de la Sociedad.

Estas reuniones permiten tomar el pulso de la filosofía que se hace en el área, los temas de mayor interés y las investigaciones más recientes. En esta ocasión se ha puesto de manifiesto la pujanza de la filosofía de la biología entre los investigadores tanto nacionales como internacionales (que pese a las dificultades derivadas de la pandemia acudieron a Salamanca). Un segundo tema de amplio interés ha sido la filosofía de la medicina, un ámbito de desarrollo relativamente reciente pero que cuenta con un importante número de investigadores con gran proyección.

De las 146 participaciones aceptadas, entre comunicaciones y simposios, la mayoría (48) se produjeron en el ámbito de la filosofía y la metodología de la ciencia, algo esperable dada la amplitud de temas que pueden desarrollarse bajo este paraguas. La segunda área de interés ha sido la filosofía de la mente y la epistemología con 25 comunicaciones, seguida de lógica, historia y filosofía de la lógica y filosofía del lenguaje que tuvieron 22 participaciones cada una. El campo de historia de la ciencia y los estudios en ciencia tecnología y sociedad contaron con 10 participaciones. Podría sorprender el escaso número de trabajos en la sección de Ciencia, Tecnología y Sociedad, pero la coincidencia con un Congreso específico del área en fechas próximas en Lisboa seguramente decantó la elección de los especialistas. Por último, la sección de argumentación tuvo 9 presentaciones, lo que se explica por el hecho de que los estudios de argumentación son un ámbito de

trabajo que se ha constituido con entidad propia, fuera de los estudios de lógica y filosofía del lenguaje, en fechas relativamente recientes.

De los 166 investigadores que acudieron al congreso un 31% fueron mujeres. Este dato, aunque escaso, mejora el porcentaje de representación de mujeres en el área, que en el curso 2019-20 constituía un 27% y este, a su vez, supone una ligera mejoría con respecto a los datos de 2014-15 donde la representación de mujeres era de un 22. En cualquier caso, la representación de mujeres en el área de Lógica y Filosofía de la Ciencia en España es una asignatura pendiente.

En esta edición se han presentado 10 simposios, lo que parece indicar una mayor preferencia de los investigadores por el trabajo colaborativo, algo no muy frecuente en la filosofía en general. Por otro lado, el contenido de los simposios refleja la heterogeneidad de intereses que se desarrollan en el área. Ana Romero de Pablos coordinó un simposio acerca de *Ontologías Híbridas*; Laura Nuño de la Rosa acerca de *Cuerpos Sexuados*; María José García Encinas sobre *Cuestiones de Metafísica Hoy*; Xavier de Donato acerca de *Deflacionismo Ontológico y Metaontológico*; Isidora Stojanovic coordinó el simposio acerca de la *Sensibilidad de Contexto de los Expresivos*, Saleta de Salvador sobre *Lenguaje y Humor*; Hubert Marraud coordinó dos simposios, uno sobre *Prácticas Argumentativas* y otro acerca de *Pragmática de las Razones*. Finalmente, Maurizio Esposito presentó uno sobre *Filosofía de la Experimentación Científica*.

Por otro lado, el Congreso aloja las *Conferencias LLullius*. Raimundus Lullius fue un filósofo medieval, nacido en Mallorca, que diseñó un método lógico para generar conocimiento. El método fue publicado en su trabajo “*Ars generalis ultima*” or “*Ars magna*” (1305), una de las primeras “ars combinatoria”. También fue una figura prominente en la época medieval, comprometido con la discusión de las ideas, la unidad de la religión y la promoción de la tolerancia. La sociedad ha elegido su nombre para estas Conferencias con la intención de representar nuestro compromiso con la búsqueda de la verdad. Los dos anteriores Conferenciantes Lullius fueron Philipp Kitcher en 2015 y Nancy Cartwright en 2018. Este año hemos contado con la participación de la filósofa estadounidense Sandra Mitchell, profesora Distinguida de la Universidad de Pittsburg y una de las principales filósofas de la biología contemporáneas.

Estamos muy satisfechas de haber celebrado este décimo congreso en el marco incomparable de la Universidad de Salamanca, más si cabe teniendo en cuenta las dificultades derivadas de la pandemia que ha impedido que muchos colegas del ámbito internacional estuvieran con nosotras.

Premio para Jóvenes investigadores e investigadoras

Evaluadas, al igual que el resto de las comunicaciones, de forma ciega por parte del comité científico, las ponencias que resultaron premiadas con el Premio para Jóvenes Investigadores e Investigadoras del X Congreso de la SLMFCE fueron:

- Laura García Portela, “Attribution methods in climate science: exploring the relevant differences”.
- Aditya Jha, “Not so distinctively mathematical explanations: microphysics matters”.
- José Antonio Pérez Escobar, “There are no “genuine mathematical explanations” in biology, but mathematical heuristics”.

2021 Lullius Lectures

Sandra D. Mitchell
Department of History and Philosophy of Science
University of Pittsburgh

From Scientific Complexity to Epistemic Pluralism

For centuries science has attempted to explain the structure and behavior of complex systems by reductive strategies: decompose the systems into the simplest components, model those in isolation, then account for the complexities by aggregating. Recent science has exposed the inadequacies of this strategy to account for many robust and dynamic complexities. I have argued for an expanded scientific epistemology of integrative pluralism in response to the demands of an ontology of complexity. This requires revisions in how we conceptualize knowledge and nature. In this talk I will present arguments in support of integrative pluralism, detailing how it transforms our understanding of scientific laws and emergence of systemic properties that cannot be fully explained by reductive strategies.

From Representational Perspectivism to Pragmatic Realism

Science acquires knowledge of nature by experience and reason. Theories and models represent what we think occurs independently of us in order to explain, predict and intervene to bring about our desired goals. I will argue that no scientific representation can be a complete mirror of nature. Every representation encodes a limited but potentially valuable perspective of nature. Partiality of representation provides epistemic, rather than ontological, reasons for pluralism. I will argue for replacing the static mirror of nature justification of realism. I will explicate and defend pragmatic realism by appeal to the interactive, provisional, coordination of conceptual resources of representation with experimental practices of causal engagement.

ARGUMENTACIÓN

Formal argumentation and epistemic logic: what can they do for each other?

Alfredo Burrieza*
University of Málaga

Antonio Yuste-Ginel**
University of Málaga

Arguing and believing are two central cognitive dimensions of both human beings and artificial intelligent agents. The interrelation of these two notions (or groups of notions) is at the root of classic debates in epistemology and argumentation theory. During this talk, we will critically review recent literature on combining two well-known families of formalisms that account respectively for argumentation and beliefs, these are, formal argumentation and epistemic logic.

Formal argumentation [2] is nowadays a well-established research area within the field of artificial intelligence that attracts the attention of scholars coming from different disciplines and traditions. It has moreover a strong potential for dealing with different theoretical and practical problems. Following [15], the history of formal argumentation can be analysed as split into two main branches: the study of *argument-based inference* and the study of *argument-based dialogues*. Regarding the former, that will be our main focus here, another broad distinction between abstract and structured approaches is usually made. *Abstract argumentation* disregards the nature of arguments and conflicts between them, treating both of them as primitive entities, and focusing on the so-called *evaluative tier* of argument strength [4], that is, how to resolve conflicts between arguments so as to sort out subsets of acceptable ones. The work of [8] is usually considered as a landmark in this sub-field. The main proposal there was to model arguments as nodes of a directed graph, and conflicts as edges between them. Although very simple, this model, called *abstract argumentation frameworks*, has been shown to be a strongly intuitive, and powerful tool for many purposes. *Structured argumentation*, on the contrary, explicitly accounts for the nature and structure of both arguments and conflicts, capturing the so-called *support* and *dialectic* tiers of argument strength [4], that focus respectively, on how strong the premisses of an argument support its conclusion and on how arguments attack

* burrieza@uma.es

** antoniyusteginel@gmail.com

and defeat each other. To this respect, [14] could be mentioned as an influential, seminal paper among structural approaches, from which the ASPIC family, and notably ASPIC⁺ [12], have received important attention.

Modern epistemic logic started with the works of [21] and [11]. Epistemic logic studies the notions of knowledge and belief (as well as related epistemic attitudes) using modal logic as the main mathematical tool. Since then, the field has evolved through the exploration of different alternatives (as non-Kripkean semantics [13]), and the solution of excessive idealizations of the model, e.g. the problem of logical omniscience. In its basic setting, to which we limit ourself here for the sake of simplicity, multi-agent epistemic models are just Kripke models (multi-directed graphs together with an atomic valuation), containing one accessibility relation R_i for each agent. Agent i believes (knows) that p is the case at a point w of the model if and only if at every R_i -successor of w , p holds as well. Epistemic models are in turn described with a multi-modal language, that adds a modality \Box_i for each agent i to the language of propositional logic, informally interpreted as “agent i believes (knows) that...”. Hence, epistemic logic can be seen as a formal model for representing higher-order, qualitative epistemic attitudes, such as “I believe that you don’t believe that I believe that p ”.

What can formal argumentation do for epistemic logic? In a nutshell, formal argumentation can provide epistemic logic with technical and conceptual insights for studying the notion of *argument-based belief*. This connection is better understood if contextualized within the relatively recent effort of the epistemic logic community to equip their models of knowledge and belief with the heretofore excluded justification component. This effort has crystallized in very heterogeneous proposals, such as justification logic [1] or neighbourhood semantics for evidence logics [20]. During the talk, we will compare semantic approaches to the notion of argument-based belief, as those developed in [18, 19] and [22], to syntactic approaches, as the one of [5, 6] (a combination of ASPIC⁺ and epistemic logic), pointing out the advantages and shortcomings of each of them.

What can epistemic logic do for formal argumentation? In few words, epistemic logic can be used as a powerful tool to deal with epistemic reasoning about abstract argumentation frameworks (and possibly other argumentation constructs). In this direction, [17] is an important antecedent, where abstract argumentation frameworks were embedded in multi-agent epistemic models under different assumptions. During the talk, we will present recent results [16, 9, 10], that establish epistemic logic as a very general framework for modelling qualitative uncertainty and multi-agency with respect to argumentation structures, as well as their dynamics, making explicit by the way what are the

epistemic assumptions underlying other existing formalisms, such as incomplete argumentation frameworks [3], and control argumentation frameworks [7].

References

- [1] Sergei Artemov and Melvin Fitting. Justification logic. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, 2016.
- [2] Pietro Baroni, Dov M Gabbay, Massimiliano Giacomin, and Leendert van der Torre. *Handbook of formal argumentation*. College Publications, 2018.
- [3] Dorothea Baumeister, Matti Järvisalo, Daniel Neugebauer, Andreas Niskanen, and Jörg Rothe. Acceptance in incomplete argumentation frameworks. *Artificial Intelligence*, 295:103470, 2021.
- [4] Mathieu Beirlaen, Jesse Heyninck, Pere Pardo, and Christian Straßer. Argument strength in formal argumentation. *IfCoLog Journal of Logics and their Applications*, 5(3):629–675, 2018.
- [5] Alfredo Burrieza and Antonio Yuste-Ginel. Basic beliefs and argument-based beliefs in awareness epistemic logic with structured arguments. In Henry Prakken, Stefano Bistarelli, Francesco Santini, and Carlo Taticchi, editors, *Proceedings of the COMMA 2020*, pages 123–134. IOS Press, 2020.
- [6] Alfredo Burrieza and Antonio Yuste-Ginel. An awareness epistemic framework for belief, argumentation and their dynamics. In Joseph Y. Halpern and Andrés Perea, editors, *Proceedings Eighteenth Conference on Theoretical Aspects of Rationality and Knowledge (TARK)*, volume 335 of *EPTCS*, pages 69–83, 2021.
- [7] Yannis Dimopoulos, Jean-Guy Mailly, and Pavlos Moraitis. Control argumentation frameworks. In Sheila A. McIlraith and Kilian Q. Weinberger, editors, *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence*. AAAI Press, 2018.
- [8] Phan Minh Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77(2):321–357, 1995.
- [9] Andreas Herzig and Antonio Yuste-Ginel. Multi-agent abstract argumentation frameworks with incomplete knowledge of attacks. In Zhi-Hua Zhou, editor, *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI-21*, pages 1922–1928. IJCAI Organization, 2021.

- [10] Andreas Herzig and Antonio Yuste-Ginel. On the Epistemic Logic of Incomplete Argumentation Frameworks. In M. Bienvenu, G. Lakemeyer, and E. Erdem, editors, *Proceedings of the 18th International Conference on Principles of Knowledge Representation and Reasoning*, pages 681–685, 11 2021.
- [11] Jaakko Hintikka. *Knowledge and belief: an introduction to the logic of the two notions*. Cornell University Press, 1962.
- [12] Sanjay Modgil and Henry Prakken. The ASPIC+ framework for structured argumentation: a tutorial. *Argument & Computation*, 5(1):31–62, 2014.
- [13] Eric Pacuit. *Neighborhood semantics for modal logic*. Springer, 2017.
- [14] John L Pollock. Defeasible reasoning. *Cognitive science*, 11(4):481–518, 1987.
- [15] Henry Prakken. Historical overview of formal argumentation. *IfCoLog Journal of Logics and their Applications*, 4(8):2183–2262, 2017.
- [16] Carlo Proietti and Antonio Yuste-Ginel. Dynamic epistemic logics for abstract argumentation. *Synthese*, 2021.
- [17] François Schwarzentruher, Srdjan Vesic, and Tjitze Rienstra. Building an epistemic logic for argumentation. In Luis Fariñas del Cerro, Andreas Herzig, and Jérôme Mengin, editors, *Logics in Artificial Intelligence*, volume 7519 of *LNCS*, pages 359–371. Springer, 2012.
- [18] C Shi, S Smets, and FR Velázquez-Quesada. Argument-based belief in topological structures. In J Lang, editor, *Proceedings of the Sixteenth Conference on Theoretical Aspects of Rationality and Knowledge (TARK)*, EPTCS. Open Publishing Association, 2017.
- [19] Chenwei Shi. No false grounds and topology of argumentation. *Journal of Logic and Computation*, 31(4):1079–1101, 2021.
- [20] Johan van Benthem, David Fernández-Duque, Eric Pacuit, et al. Evidence logic: A new look at neighborhood structures. volume 9 of *Advances in modal logic*, pages 97–118. College Publications, 2012.
- [21] George H Von Wright. *An essay in modal logic*. 1953.
- [22] Yi N Wáng and Xu Li. A logic of knowledge based on abstract arguments. *Journal of Logic and Computation*, 2021.

Obtención de explicaciones orientadas al humano a partir de demostraciones automáticas de teoremas: un algoritmo de transformación

Assumpció Casanova*
Universitat Politècnica de València

En esta comunicación se presenta un sistema de inferencia que permite deducir las mismas cláusulas que un demostrador automático de teoremas sobre un formato clausular. Además, se presenta un algoritmo que transforma las secuencias de pasos realizada por el demostrador como una deducción en el sistema formal, con lo que se produce una vista orientada al humano.

La mecanización de la demostración de teoremas iniciada a finales de los 50 del siglo pasado dio lugar a dos líneas de investigación divergentes en objetivos y métodos, tal y como se indica en (Davis, 2001). La primera de ellas pretendía la elaboración de un razonamiento lógico matemático similar al humano, combinando sistemas formales de deducción y heurísticos. La segunda línea contemplaba la demostración automática como un objetivo en sí mismo, prefiriendo la eficiencia a la inteligibilidad de la prueba, y de ahí que se denomine “orientados a la máquina” a los métodos de demostración enmarcados en esta línea.

Los más conocidos de estos últimos aparecieron en la década de los 60, destacando el método de *resolución* de Robinson y el de *eliminación de modelos* (EM) de Loveland, definidos para el formato clausular (Bibel, Eder, 1993). En el lenguaje de primer orden, una demostración consiste en encontrar, mediante el algoritmo de *unificación de términos*, un conjunto inconsistente de instancias en el *Universo de Herbrand* del conjunto de cláusulas obtenido a partir de las fórmulas de partida, junto con la negación del teorema a demostrar. El lenguaje se limita al cuantificador universal, eliminando los existenciales mediante la introducción de *funciones de Skolem*, lo que preserva la inconsistencia.

Las décadas posteriores se centraron en desarrollar estrategias y refinamientos que priorizaron una implementación eficiente de estos métodos. Ello se tradujo en su utilización exitosa en ciertos ámbitos de demostración de teoremas, y en otras aplicaciones propias de la Inteligencia Artificial. Pero medio siglo después, aún se están proponiendo nuevas interfaces cuyo propósito es ayudar a la comprensión de una demostración por resolución y sus variantes, como por ejemplo, la descrita en (Heltz, Libal, Riener y Rukhaia, 2013).

Otra aplicación notable del método de resolución se dio en los lenguajes de *Programación Lógica* (PL), como Prolog, en donde un programa viene dado por un conjunto de

* casanova@dsic.upv.es

cláusulas de Horn definidas, cuantificadas universalmente. Desde los 70 existen implementaciones eficientes de dicho lenguaje basadas en resolución SLD (*Selective Linear Definite Clause Resolution*). La inclusión en el lenguaje del símbolo de implicación permite obtener una expresión más natural de las cláusulas definidas, y junto a la estrategia SLD, facilita leer los cálculos realizados por el programa como una secuencia de pasos de inferencia (*instanciación y modus ponens*), lo que ha influido sin duda en su amplia difusión en diversos campos.

En los 80 se desarrolló la *Programación Lógica Disyuntiva* (PLD), extensión de la PL a cláusulas generales, presentada con detalle en (Lobo, Minker, Rajasekar, 1992). Desde entonces, y hasta ya entrados en el presente siglo, se han venido realizando diversas implementaciones, basadas en refinamientos de resolución y de EM, aunque la disyunción dificulta una interpretación procedural al estilo de Prolog. Este es uno de los ejemplos de uso de cláusulas disyuntivas que motiva la propuesta que se presenta en esta comunicación.

Dicha propuesta consiste en una interfaz cuyo objetivo es favorecer la interpretación de demostraciones automáticas realizadas, en concreto, por un demostrador EM.

En primer lugar, se propone un sistema de inferencia para el formato clausular con las reglas de inferencia *conmutatividad e idempotencia de literales, contraposición, instanciación y una regla de corte* o adaptación al formato clausular de una generalización de *modus ponens*:

$$\frac{\forall \mathbf{x}(L_1(\mathbf{x}) \vee D_1(\mathbf{x})), \dots, \forall \mathbf{x}(L_n(\mathbf{x}) \vee D_n(\mathbf{x})) \quad \forall \mathbf{x}(L_1(\mathbf{x}) \wedge \dots \wedge L_n(\mathbf{x}) \rightarrow D(\mathbf{x}))}{\forall \mathbf{x}(D_1(\mathbf{x}) \vee \dots \vee D_n(\mathbf{x}) \vee D(\mathbf{x}))}$$

siendo \mathbf{x} una tupla de variables, $L_i(\mathbf{x})$ un literal y $D_i(\mathbf{x})$ una disyunción de literales.

En esta regla adaptada se sintetizan diferentes pasos realizados a nivel del demostrador automático, como la instanciación a términos nuevos introducidos por *skolemización*, que pueden ser generalizados tras aplicaciones de diferentes pasos de resolución de literales. La regla de *modus ponens* adaptada es el caso particular en que todos los $D_i(\mathbf{x})$ son vacíos, y se mantiene en el sistema por claridad de las trazas.

Todas estas reglas son universalmente válidas, por lo que se trata de un sistema correcto.

El sistema es completo para ciertas subcláusulas de las consecuencias lógicas, entendiéndose que $C' = \forall \mathbf{x}D'(\mathbf{x})$ es una subcláusula de $C = \forall \mathbf{x}D(\mathbf{x})$ si todos los literales de $D'(\mathbf{x})$ aparecen en $D(\mathbf{x})$. Así, sea Γ un conjunto no contradictorio de cláusulas universales, y C una cláusula no tautológica tal que $\Gamma \models C$. Se dice que C' subcláusula de C es una *simplificación de C bajo Γ* sii $\Gamma \models C'$ y ninguna de las subcláusulas propias de C' es consecuencia lógica de Γ . Se denota C_Γ al conjunto de todas las cláusulas que son una simplificación de C bajo Γ . Entonces se tiene que

$$\Gamma \models C \text{ sii } \forall C' \in C_\Gamma, \Gamma \vdash C'$$

Con lo que el sistema de inferencia propuesto es completo para la simplificación de las cláusulas.

En segundo lugar, se expone un algoritmo que transforma las trazas de los cálculos realizados por EM para demostrar una cláusula, en demostraciones en este sistema de inferencia. Se obtiene como resultado una cadena deductiva que termina en la cláusula a demostrar.

Este algoritmo lee los pasos dados por el demostrador EM como aplicación de las reglas del sistema formal, que infiere cláusulas a partir de cláusulas, generándose una traza de $\Gamma \vdash C$ a partir de la demostración EM de inconsistencia de Γ y $\neg C$. Con ello se consigue una presentación de las trazas de ejecución del demostrador en un lenguaje formal “orientado al humano”.

Se presentan ejemplos tomados de diferentes ámbitos, como el de *razonamiento por excepciones y resolución de problemas lógicos* (Lusk, Overbeek, 1985).

Finalmente, se discute el mecanismo de *extracción de respuestas* propio de la PL y de la PLD, y su integración en el sistema formal propuesto. En el paradigma PLD, un programa viene dado por un conjunto Γ de cláusulas al que se le proponen preguntas de la forma $\exists \mathbf{x}(L_1(\mathbf{x}) \wedge \dots \wedge L_k(\mathbf{x}))$.

En el caso en que $\Gamma \models \exists \mathbf{x}(L_1(\mathbf{x}) \wedge \dots \wedge L_k(\mathbf{x}))$, una respuesta a la pregunta es una fórmula $\varphi = \text{Respuesta}(\mathbf{t}_1/\mathbf{x}) \vee \dots \vee \text{Respuesta}(\mathbf{t}_m/\mathbf{x})$, tal que $\Gamma \models \varphi$, en donde *Respuesta*(\mathbf{x}) denota la conjunción $L_1(\mathbf{x}) \wedge \dots \wedge L_k(\mathbf{x})$. Dado que las cláusulas de Γ son universales, el teorema de Herbrand asegura en tal caso la existencia de una respuesta de la forma anterior.

Es habitual presentar o referirse a la respuesta como el conjunto de sustituciones $\{\mathbf{t}_1/\mathbf{x}, \dots, \mathbf{t}_m/\mathbf{x}\}$, aunque ello puede suponer un debilitamiento del resultado. En el demostrador EM así como el sistema formal de la interfaz, en los que se producen cláusulas a partir de cláusulas, se genera la respuesta a partir del conjunto de cláusulas de la forma dual equivalente, o sus formas simplificadas. Estas últimas darían un resultado en general lógicamente más fuerte que la mera respuesta.

Por ejemplo, $(P(a) \wedge Q(a)) \vee (P(b) \wedge Q(b))$ tiene la forma de una respuesta a la pregunta $\exists x(P(x) \wedge Q(x))$. Un demostrador podría obtener esta respuesta a partir de un programa tal que $\Gamma \models P(a) \vee P(b)$, $P(a) \vee Q(b)$, $Q(a) \vee P(b)$, $Q(a) \vee Q(b)$, cláusulas equivalentes lógicamente a la respuesta, o también de un programa tal que $\Gamma \models P(a)$, $Q(b)$, $Q(a) \vee P(b)$, cláusulas simplificadas que implican las anteriores, y que por lo tanto proporcionan una información más precisa.

Existen diferentes vías de incorporar la extracción de respuestas en el sistema formal. Algunas precisan de ciertas extensiones que permitan al algoritmo de transformación estructurar la traza de la demostración para presentar la respuesta como una conclusión, usando hipótesis auxiliares y el teorema de la deducción. Otras son más directas como, por ejemplo, la que consiste en añadir a Γ una cláusula universal que forma parte de la definición explícita de un nuevo predicado *Respuesta*(\mathbf{x}), a partir de los literales de la pregunta. Esta definición no precisa de más justificación, por tratarse de una extensión

conservativa del programa de partida. El uso por parte del demostrador EM de diferentes instancias de esta cláusula conduce directamente en la correspondiente traza a una respuesta clausular disyuntiva, tan general como sean capaces de obtener los refinamientos y podas del demostrador automático. No obstante, la traza en el sistema formal visualiza y remarca las cláusulas simplificadas de la forma dual de la respuesta que se hubieran generado en la demostración automática, información que aparece más debilitada en la respuesta final.

Referencias

- [1] Bibel, W., Eder, E. (1993). “Methods and Calculi for Deduction”. En D. M. Gabbay, C. J. Hogger y J. A. Robinson (eds.) *Handbook of Logic in Artificial Intelligence and Logic programming*. Vol. 1. Clarendon Press, Oxford, pp. 67-182.
- [2] M. Davis. (2001). The Early History of Automated Reasoning. *Handbook of Automated Reasoning*. Vol. I. (A. Robinson, A. Voronkov, eds.), Elsevier Science, pp. 3-15.
- [3] S. Hetzl, T. Libal, M. Riener, M. Rukhaia. (2013). “Understanding Resolution Proofs through Herbrand’s Theorem”. En D. Galmiche y D. Larchey-Wendling (eds.). *TABLEAUX 2013*, LNCS 8123: 157-171. Springer-Verlag, Berlin.
- [4] E. Lusk, R. Overbeek (1985). “Non-Horn Problems”. *Journal of Automated Reasoning*, 1:103-114.
- [5] Lobo, J., Minker, J., Rajasekar, A. (1992). *Foundations of Disjunctive Logic Programming*. The MIT Press, Cambridge, MA.

The interactionist approach to speech acts and its role in the analysis of argumentation

Amalia Haro Marchal*
University of Granada

Among the different argumentation theories that adopt speech act theory as their theoretical framework, the Pragma-dialectics model proposed by Frans H. Van Eemeren and Rob Grootendorst (1984) and the Linguistic Normative Model of Argumentation developed by Lilian Bermejo-Luque (2011) are two of the most influential proposals. Both theories provide different models for the analysis of examples of argumentation. Speech acts theories applied by these proposals and which have their origin in John L. Austin's theory (1962) are framed within the so-called "institutionalist" (Searle 1969) and "intentionalist" (Bach and Harnish 1979) theories.

The main goal of this presentation is to identify what counts as argumentation according to the models proposed by these theories in order to show that, if the speech act theory they use is replaced by an interactionist theory (Clark 2006, Sbisà 2006, Carassa and Colombetti 2009), the analysis of the examples of argumentation and the unity of the argumentative discourse itself differs. One advantage of the application of the interactionist theory to the models mentioned above is that it allows us to effectively explain argumentation as a communicative practice, in which both speaker and hearer play a fundamental role.

The Pragma-dialectics Model (PDM) conceives the argumentation as a complex illocutionary act formed by elementary illocutions which have the illocutionary force of assertions. These illocutions constitute an illocutionary set that stands in a relation of justification or refutation of an expressed opinion that functions as a premise or conclusion but that is not part of the speech act of arguing. The speech acts carried out in the argumentative discourse present two simultaneous illocutionary forces: at the sentential level, the speech acts belong to the category of assertives; at the textual level, they have the illocutionary force of argumentation.

In the Linguistic Normative Model of Argumentation (LNMA), argumentation is characterized as a single speech act composed of two second-order speech acts: the act of adducing (a reason) and the act of concluding (a target-claim). These speech acts are second-order constative speech acts that result from a certain relation between two first-order speech acts, a relation that is specified in the implicit inference-claim between the contents of these first-order speech acts.

* amaliaharomarchal@gmail.com

In order to see how these models can be applied to the analysis of examples of argumentation, I will analyze a dialogue from a scene in the movie “When Harry met Sally”. This dialogue is an example that Horn (2004: 5) analyzed concerning the difference between conversational and generalized implicatures. In this presentation, it will be used as an example of argumentation. Harry is talking about his woman friend Sally with his friend Jess:

- (1) a. [Jess:] If she’s so great, why aren’t YOU taking her out?
- b. [Harry:] How many times do I have to tell you, we’re just friends.
- c. **[Jess:] So you’re saying she’s not that attractive.**
- d. [Harry:] No, I told you she is attractive.
- e. **[Jess:] But you also said she has a good personality.**
- f. **[Harry:] She has a good personality.**
- g. [Jess:] [Stops walking, turns around, throws up hands, as if to say “Aha!”]
- h. [Harry:] What?
- i. **[JESS:] When someone’s not that attractive they’re ALWAYS described as having a good personality.**
- j. **[HARRY:] Look, if you were to ask me what does she look like and I said she has a good personality, that means she’s not attractive. But just because I happen to mention that she has a good personality, she could be either. She could be attractive with a good personality or not attractive with a good personality.**
- k. [Jess:] So which one is she?
- l. [Harry:] Attractive.
- m. [Jess:] But not beautiful, right?

From PDM, the analysis of the example would be as follows:

- By means of (1c), Jess carries out a speech act of concluding that does not form part of the argumentation. The content of (1c) is the expressed opinion to which the argumentation is related.
- At the sentential level, the argumentation is formed by an assertive speech act that Jess carries out by means of (1e) and by an unexpressed premise that makes (1e) an argumentation at the textual level.
- The elementary illocution (1e) together with the unexpressed premise constitute a speech act of arguing because they are in a relationship of justification with the expressed opinion in (1c).

From LNMA, the analysis of the example would be as follows:

- By means of (1c) Jess asserts that Harry meant that Sally is not very attractive.

- By means of (1e) Jess asserts that Harry told him that she had a good personality.
- Jess implicitly asserts that, if Harry told him that she had a good personality, then Harry meant that she's not very attractive.
- Jess adduces that Harry said that she had a good personality as a reason to - conclude that he meant her to be unattractive.
- The argumentation, according to LNMA, would be constituted by (1c) and (1e) together with the implicit assertion.

As we shall now see, by applying the interactionist speech act theory to the PDM and LNMA models, rather than the institutionalist and intentionalist approaches they apply, the subject matter would change; as a consequence, what is considered as a case of argumentation under both models would cease to be so.

If the interactionist speech act theory is applied to PDM, the example would be analyzed as follows:

- In order to know if Jess's speech act constitutes an argumentation, we must take into account Harry's response. By means of (1f), Harry shows that he does not accept that Jess's utterance (1e) is in a relationship of justification with the opinion expressed by means of (1c).
- Insofar as it functions as what makes explicit the unexpressed premise, (1i) is needed for his speech act to constitute an argumentation. By (1i), Jess shows that he does not accept Harry's interpretation of (1e) as a mere assertion.
- By (1j), Harry shows that he accepts Jess' speech act as an argumentation, but he questions the warrant.
- Thus, the argumentation constituted by (1e) along with the warrant would cease to be one.

This analysis shows that, by applying the interactionist speech act theory to PDM, the argumentation constituted by (1e) together with the warrant would cease to be an argumentation.

Applying the interactionist theory of speech acts to LNMA, the analysis of the example would be as follows:

- At least Harry's response (1f) is needed for (1c), (1e) and the implicit assertion to constitute an argumentation.
- However, by means of (1f), Harry shows that he does not consider Jess's utterances as a speech act of arguing.
- Jess, by means of (1i), makes explicit what supports his implicit assertion, rejecting Harry's interpretation of his speech acts as mere assertions.
- By means of (1j), Harry would have accepted that Jess's speech acts constitute an argumentation, but he calls into question the warrant.

- Harry shows that Jess’s argumentation, even if it is a case of argumentation, is bad.

The analysis shows that, as in PDM, an interactionist speech act theory applied to LNMA results in a modification of the unity of the argumentative discourse. Thus, the argumentation constituted by (1c) and (1e) would in fact no longer be a case of argumentation.

As we have seen, the analysis of the dialogue has shown that there could be argumentations in which what a speaker means does not coincide with the meaning jointly construed by the speaker and the hearer. And these cases can be explained if the interactionist speech act theory is applied to models like PDM and LNMA.

References

- [1] Austin, J.L. (1962). *How to Do Things with Words*. Oxford: Oxford University Press.
- [2] Bach, K. & Harnish, R.M. (1979). *Linguistic Communication and Speech Acts*. Cambridge: Mass, MIT Press.
- [3] Bermejo-Luque, L. (2011). *Giving Reasons: A Linguistic-Pragmatic Approach to Argumentation Theory*. Dordrecht: Springer.
- [4] Carassa, A. & M. Colombetti (2009). “Joint Meaning”. *Journal of Pragmatics*, 41, 1837-1854.
- [5] Clark, H. H. (2006). “Social Actions, Social Commitments”, in S. C. Levinson & N. J. Enfield (eds.) *Roots of Human Sociality: Culture, Cognition, and Human Interaction*. Oxford: Berg Press, 126-150.
- [6] van Eemeren, F.H., & Grootendorst, R. (1984). *Speech-acts in argumentative discussions. A theoretical model for the analysis of discussions directed towards solving conflicts of opinion*. Foris, Dordrech/Mounton de Gruyter: Berlin.
- [7] Horn, L. R. (2004). “Implicature”, in L. R. Horn & G. Ward (eds.) *The Handbook of Pragmatics*. Oxford: Blackwell Publishing, 3-28.
- [8] Sbisà, M. (2006). “Communicating Citizenship in Verbal Interaction: Principles of a Speech Act Oriented Discourse Analysis”, in H. Hausendorf & A. Bora (eds.). *Analyzing Citizenship Talk*. Amsterdam: John Benjamins Publishing Company, 151-180.
- [9] Searle, J.R. (1969). *Speech Acts: an Essay in the Philosophy of Language*. Cambridge, Mass: Cambridge University Press.

Challenging authority with argumentation: The pragmatics of not accommodating authority

Marcin Lewinski*

ArgLab, Nova Institute of Philosophy, Universidade Nova de Lisboa

The goal of this contribution is to propose a new account of challenging authority in argumentative discourse. To this end, I draw on the resources of recent speech act theory and argumentation theory. As early as in his *Rhetoric*, Aristotle (2007) claimed that the ethos, or character of the speaker, is “the most authoritative form of persuasion.” In turn, contemporary speech act theory demonstrates how for virtually all speech acts to be felicitously performed, the speaker needs to be in a position of an epistemic (theoretical) or deontic (practical) authority (Austin, 1962; Searle, 2010). But while fundamental in understanding how authority functions in discourse, these classic approaches have a specific, and somewhat limited, focus: Aristotle looked exclusively into the speaker’s authority established explicitly in discourse, whereas speech act theorists typically draw on pre-established and formally recognised “deontic powers” of speakers.

Recent work in speech act theory, however, points to other, more sophisticated forms of discourse-related authority, namely, to informal authority gained by default via presupposition accommodation (Bianchi, 2019; Langton, 2015; 2018a; 2018b; Witek, 2013; 2019). Inspired by insights from Austin (1962) and Lewis (1979) this work highlights how authority can be established on the fly as conversation develops. A group member who says, “Alright guys, where do we go for dinner?!” can establish herself as a group leader, should other members go along with her taking the initiative. But there are much more troublesome examples of this phenomenon, especially those related to all forms of discriminatory speech. A white male passenger who on a subway addresses an Arab woman with “F***in’ terrorist, go home. We don’t need your kind here” acquires a certain practical authority as a guardian of the alleged rules of the society (Maitra, 2012; Langton, 2018b). This authority can be accommodated as such should nobody – the addressed hearer or other bystanders – challenge the speaker’s entitlement to enforce the rules that don’t exist. In this context, various strategies of “defying,” “blocking,” “undoing,” or otherwise challenging the purported authority are of key importance (Kukla, 2014; Lance & Kukla, 2013; Langton, 2018a; Caponetto, 2020).

In this paper, I extend this work on challenging authority by treating such conversational phenomena as argumentative phenomena. I first analyse such presupposed claims to authority as “virtual standpoints” of speakers-qua-protagonists (Jackson, 1992; van

* m.lewinski@fcsh.unl.pt

Eemeren et al., 1993; Lewiński & Oswald, 2013). When challenged or “called out” by hearers-qua-antagonists these standpoints open the broader “disagreement space” of argumentative discourse. Such challenges characteristically consist in more or less explicit critical questions and counter-arguments to the speaker’s “virtual” argument from authority (“Who are you to tell..?!?” “You have no right to tell...”). Here, the extensive work on the argument from authority – and the fallacious argumentum ad verecundiam – provides a rich reservoir of methods (Goodwin, 1998; Koszowy & Walton, 2019; Wagemans, 2011; Willard, 1990). Since both epistemic and deontic authority eventually depend on social recognition, the basic dynamics identified in this study rely on explicitly or implicitly denying this recognition to the abusive speaker. Importantly, as shown through a number of actual examples taken from online exchanges, these argumentative and pragmatic dynamics can involve participants other than the addressed hearers themselves (Lewiński, 2021).

In this way, the paper shows how nuanced yet fundamental problems of public discourse identified by philosophers of language can become an inspiration and an object of study for argumentation theorists. It also shows how argumentation theory can feed back into a broad pragmatic study of discourse thanks to its developed framework of concepts and methods for understanding and evaluating forms of discourse where disagreements are managed.

References

- [1] Aristotle. (2007). *On rhetoric: A theory of civic discourse* (Kennedy, G.A., trans.), 2nd edition. Oxford: Oxford University Press.
- [2] Austin, J.L. (1962). *How to do things with words*. Oxford: Clarendon Press.
- [3] Bianchi, C. (2019). Asymmetrical conversations: Acts of subordination and the authority problem. *Grazer Philosophische Studien*, 96(3), 401–418.
- [4] Caponetto, L. (2020). Undoing things with words. *Synthese*, 197(6), 2399–2414.
- [5] Eemeren, F.H. van, Grootendorst, R., Jackson, S., & Jacobs, S. (1993). *Reconstructing argumentative discourse*. Tuscaloosa: University of Alabama Press.
- [6] Goodwin, J. 1998. Forms of authority and the real ad verecundiam. *Argumentation*, 12(2), 267–280.
- [7] Jackson, S. (1992). “Virtual standpoints” and the pragmatics of conversational argument. In F.H. van Eemeren, R. Grootendorst, J.A. Blair & C.A. Willard (Eds.), *Argumentation illuminated* (pp. 260-269). Amsterdam: SicSat.

- [8] Kukla, R. (2014). Performative force, convention, and discursive injustice. *Hypatia*, 29(2), 440–457.
- [9] Lance, M., & Kukla, R. (2013). ‘Leave the gun; take the cannoli’: The pragmatic topography of second-person calls. *Ethics*, 123(3), 456–478.
- [10] Langton, R. (2015). How to get a norm from a speech act. *The Amherst Lecture in Philosophy*, 10, 1–33.
- [11] Langton, R. (2018a). Blocking as counter-speech. In D. Fogal, D.W. Harris, & M. Moss (eds.), *New work on speech acts* (pp. 144-164). Oxford: Oxford University Press.
- [12] Langton, R. (2018b). The authority of hate speech. In J. Gardner, L. Green, & B. Leiter (eds.), *Oxford Studies in Philosophy of Law*, Volume 3 (pp. 123–152). Oxford: Oxford University Press.
- [13] Lewiński, M. (2021). Illocutionary pluralism. *Synthese*, DOI: <https://doi.org/10.1007/s11229-021-03087-7>.
- [14] Lewiński, M. & Oswald, S. (2013). When and how do we deal with straw men? A normative and cognitive pragmatic account. *Journal of Pragmatics*, 59, 164-177.
- [15] Lewis, D. (1979). Scorekeeping in a language game. *Journal of Philosophical Logic*, 8(1), 339–359.
- [16] Maitra, I. (2012). Subordinating speech. In I. Maitra & M. K. McGowan (eds.), *Speech and harm: Controversies over free speech* (pp. 94-120). Oxford: Oxford University Press.
- [17] Koszowy, M., & Walton, D. (2019). Epistemic and deontic authority in the argumentum ad verecundiam. *Pragmatics and Society*, 10(2), 151-179.
- [18] Searle, J.R. (2010). *Making the social world: The structure of human civilization*. Oxford: Oxford University Press.
- [19] Wagemans, J.H.M. (2011). The assessment of argumentation from expert opinion. *Argumentation*, 25(3), 329-339.
- [20] Willard, C.A. 1990. Authority. *Informal Logic*, 12, 11-22.
- [21] Witek, M. (2013). How to establish authority with words: Imperative utterances and presupposition accommodation. In A. Brozek, J. Jadacki, & B. Žarnic (Eds.), *Theory of imperatives from different points of view* (2) (pp. 145-157). Warszawa: Semper.
- [22] Witek, M. (2019). Illocution and accommodation in the functioning of presumptions. *Synthese*, DOI: <https://doi.org/10.1007/s11229-019-02459-4>.

Un argumento falaz para defender el multiverso frente al diseño inteligente

Fernando Migura*
UPV/EHU

Lo improbable de un universo sintonizado tan afinadamente como el nuestro, que hace posible no solo la vida sino nuestra presencia como observadores, requiere una explicación no teleológica (teológica o sobrenatural). El recurso al argumento probabilístico (inferencia bayesiana) ha orientado la búsqueda hacia la hipótesis del multiverso. Pero como trataré de mostrar, este recurso es falaz pues se construye sobre una doble confusión entre sucesos-tipo y sucesos ejemplar y entre probabilidades a priori y probabilidades a posteriori. Nuestro punto de partida es la que Hacking bautiza como “Falacia del jugador inversa” (“The Inverse Gambler’s Fallacy” (Hacking (1987)), una versión retrospectiva de la conocida falacia del jugador (Gambler’s Fallacy, (Tversky and Kahneman (1974))).

Como sabemos, cualquiera que juegue a los dados comete la falacia del jugador cuando confunde acontecimientos lógicamente independientes con los que son lógicamente dependientes. La jugadora está dispuesta a apostar a favor de que la siguiente tirada será un doble-seis, tras 35 tiradas en las que no se ha obtenido un doble-seis porque sabe que la probabilidad de obtener un doble-seis en 36 tiradas es $2/3$. Quien así razona aplica mal el principio según el cual la probabilidad de un acontecimiento raro o improbable aumenta con el número de ensayos y no distingue entre la probabilidad de una sucesión de lanzamientos y el último lanzamiento de la sucesión o, si los distingue, no los toma como acontecimientos lógicamente independientes.

1 La falacia de la jugadora inversa

Hacking (Hacking (1987)) nos muestra que la jugadora también razona mal retrospectivamente, cuando para explicar un acontecimiento improbable, como, p. e., el doble-seis obtenido por otro jugador que presencia al acercarse a la mesa de juego en el casino; supone que ese doble-seis es el resultado de un gran número de lanzamientos previos. Hacking además insiste en que comete un error el habitante de nuestro mundo sintonizado que aplica un razonamiento análogo para, ante lo improbable de nuestro mundo, inferir que probablemente es el último de una larga sucesión fluctuante de universos generados desde el Big Bang que van instanciando sucesivamente de manera lógicamente

* ylpimizaf@ehu.eus

independiente todos los valores posibles para los parámetros físicos involucrados (esto es, de un multiverso en la versión del modelo cosmológico de Wheeler (Wheeler (1977)) para así, eludir verse forzado a defender un mundo diseñado ex profeso por el Creador para albergarnos.

Así, para Hacking apostar por una pluralidad de universos para explicar la especificidad del nuestro es falaz. Nosotros coincidimos en esto con Hacking, es decir, en que hay involucrada una falacia, pero además damos una explicación completa de la misma que justifica tal calificación ante las objeciones y contraargumentos propuestos por los que tratan de rehabilitar dicho argumento defendiendo su corrección con cualificaciones de distinto tipo como la apelación al principio antrópico (Carter (1974)) entendido como efecto de elección de la observación (entre otros, Leslie (1988), McGrath (1988), Bromtom (2002), Landsman (2016), Friederich (2018))

Así, si identificamos nuestro-improbable-mundo-sintonizado-que permite-nuestra-existencia-como-observadores con otro suceso improbable como es la obtención de un doble-seis jugando a los dados; sabemos que cuantas más veces lancemos los dados, más probable es que obtengamos al menos un doble seis en la sucesión de lanzamientos. Atendiendo a los principios básicos de la probabilidad sabemos que la probabilidad de obtener un resultado aumenta con la repetición de los ensayos. Pero independientemente del número de ensayos, también sabemos que la probabilidad de que obtengamos un doble-seis es de $1/36$ para cada tirada. Como hemos visto, incurre en la conocida falacia del jugador quien tras 35 intentos está dispuesta a apostar de manera desmesurada a favor de un doble-seis en el siguiente lanzamiento. Incurre en la falacia retrospectiva del jugador quien al acudir al casino y presenciar un doble-seis infiere que probablemente es el último lanzamiento de una sucesión más o menos larga de ensayos previos.

La falacia en cosmología se construye sobre la suposición de que un mundo improbable como el nuestro proporciona información relevante para inferir una pluralidad de universos previa. Significa razonar retrospectivamente suponiendo que esa pluralidad de ensayos previos sirve para explicar ese resultado que sería más improbable en un único ensayo y, además, asumiendo que esta es la mejor explicación disponible.

2 La jugadora confunde la probabilidad a priori con la probabilidad a posteriori

Aunque Hacking no lo formula en estos términos, se da una ambigüedad tipo/ejemplar en el uso de la expresión “lograr un doble-seis”. Si se interpreta como tipo estamos hablando a priori de algún doble seis: del primero que se da o del último (si la sucesión de lanzamientos termina al lograr uno). Sabemos que cuantos más ensayos se realicen es más probable que se dé algún resultado favorable (predicción de un tipo de suceso). Por tanto, al encontrarnos ante un (algún) resultado favorable: “el doble-seis” con el que se termina el juego, tendemos a pensar que para obtenerlo la sucesión de ensayos ha sido más bien larga que corta (explicación de un resultado específico). Si optamos por

entender “lograr un doble-seis” en su versión más específica como ejemplar: como “este doble seis” ante el que me encuentro, reparamos en el error al razonar. En última instancia, oscilamos entre un razonamiento que nos empuja a la falacia y otro que no, porque el resultado (acontecimiento) caracterizado a priori con una descripción definida (p. e., “el primer doble-seis logrado”) alude al tipo (a algún doble-seis) y no coincide exactamente con el verdadero ejemplar: la instancia del tipo señalada a posteriori con un demostrativo (directamente referencial) como “este” (en “este doble-seis”). De manera análoga en su interpretación cosmológica, si hablamos de nuestro universo, hablamos de este en el que yo existo, no de otro cualquiera en que pudieran darse otros agentes cognoscentes. La pluralidad de universos aumenta la probabilidad de universos como el nuestro pero no de este nuestro que es altamente improbable (siempre suponiendo que es lógicamente independiente de los demás). Cuando hablamos del mundo que habitamos nosotros, no hablamos de tipos de universos sino de este-nuestro-universo (U) dando al demostrativo este y al deíctico nuestro la lectura directamente referencial que les atribuye Kaplan (Kaplan (1989)) en su lógica de los demostrativos. Yo soy fruto de este universo nuestro sutilmente sintonizado y el pronombre personal de primera persona yo opera como deíctico puro instanciado (Kaplan (1989)) o como designador rígido, si se prefiere la terminología de Kripke (Kripke (1980)). Yo existo aquí y ahora en este universo y no podría existir en otro si aceptamos el esencialismo de origen característico de la teoría kripkeana, por lo que necesariamente existo en este nuestro universo y no podría existir en otro. Obviamente, esto no significa que mi existencia o la de nuestro universo sea necesaria en ningún sentido.

Conclusión

Cuando decimos atendiendo a los principios básicos de la probabilidad y la ley de los grandes números que un mayor número de acontecimientos azarosos aumentan la probabilidad de un fenómeno extraño/improbable, estamos aludiendo a tipos de acontecimientos a priori y la predicción es razonable. Pero como hemos visto, cuando tratamos de utilizar dicha ley en la explicación de una instancia dada a posteriori de un suceso improbable, infiriendo que probablemente es fruto de un número elevado de ensayos previos, incurrimos en la falacia que denuncia Hacking. Ahora bien, si tuviéramos que explicar cómo pueden darse universos lo suficientemente sintonizados para albergar observadores, en el sentido de sujetos cognoscentes de algún tipo, la hipótesis de la pluralidad de universos que se van instanciando sería perfectamente razonable.

Referencias bibliográficas

- [1] Bostrom, N. (2002), *Anthropic Bias: Observation Selection Effects in Science and Philosophy*. Nueva York: Routledge.
- [2] Carter, B. (1974), Large number coincidences and the anthropic principle in cosmology. In: *Confrontation of cosmological theories with observational data; Proceedings*

- of the Symposium*, Krakow, Poland, September 10-12, 1973. (A75-21826 08-90) Dordrecht, D. Reidel Publishing Co., 1974, p. 291-298.
- [3] Friederich, S. (2018) Reconsidering the inverse gambler's fallacy charge against the fine-tuning argument for the multiverse. *Journal for General Philosophy of Science*, 50. pp. 29-41.
- [4] Hacking, I. (1987), The Inverse Gambler's Fallacy: the Argument from Design. The Anthropic Principle Applied to Wheeler Universe. *Mind*, Vol. XCVI, (383):331–340.
- [5] Kaplan, D. (1989), Demonstratives. In J. Almog, J. Perry and H. Wettstein (eds) *Themes from Kaplan*. New York: Oxford University Press.
- [6] Kripke, S. (1980), *Naming and Necessity*. Cambridge Mass.: Harvard University Press.
- [7] Landsman, K. (2016), The Fine-Tuning Argument: Exploring the Improbability of Our Existence. In Landsman, K. and E. van Wolde (eds) *The Challenge of Chance* (The Frontiers Collection), SpringerOpen, pp. 11-130.
- [8] Leslie, J. (1988), No Inverse Gambler's Fallacy in Cosmology. *Mind* 97 (386):269-272.
- [9] McGrath, P. J. (1988), The Inverse Gambler's Fallacy and Cosmology– A Reply to Hacking. *Mind* 97 (386): 265-268.
- [10] Tversky, A and D. Kahneman (1974), Judgment under Uncertainty: Heuristics and Biases. *Science* 27 Sep 1974: 185, (4157): 1124-1131.
- [11] Wheeler, J. A (1977), Genesis and Observership. In R. E. Butts and J. Hintikka (eds) *Foundational Problems in Special Sciences. Part Two of the Proceedings of the Fifth International Congress of Logic, Methodology and Philosophy of Science*, London, Ontario, Canada-1975, pp. 3-33.

De la argumentación en espacios emergentes, una propuesta entre el juego y la práctica semántica

Diana Lizbeth Ruiz Rincón*
Universidad Autónoma de Chiapas

Francisco Javier González Rivas**
Universidad Autónoma de Chiapas

En el inicio todo era juego o, al menos susceptible de convertirse en ello. Pero para precisar el universo de discurso, debemos definir el tipo de juego al que nos estamos refiriendo, cuidando así de dejar fuera a aquellas interacciones que se encuentren motivadas por la emotividad, el azar o la irracionalidad. De los tipos de interacción de la que nos ocuparemos serán aquellos a los que podemos atribuir un comportamiento racional entre los agentes involucrados. Morgenstern, en su artículo intitulado “La teoría de los Juegos y del comportamiento económico” (1955) relaciona al comportamiento racional, con el más adecuado pues, dejando de lado el contexto disciplinar en el que nace la Teoría de Juegos, resulta un modelo que se centra sí, en la interacción, pero también en su resultado, focalmente en las disposiciones epistémicas y estratégicas empleadas por cada uno de los agentes participan.

Aspectos reveladores del modelo que nos ofrece la Teoría de Juegos (TJ), son: i) describe un específico contexto lingüístico; ii) define a la elección, esto es, que un conjunto de reglas estructurales que plantean y delimitan aquellas decisiones que pueden ser tomadas; iii) estructura, organiza y representa el conjunto de decisiones que en cada interacción van tomando los agentes epistémicos, y; iv) a partir tanto de las reglas establecidas, como del desarrollo del juego, es posible para cada jugador o agente epistémico, diseñar una estrategia que le permita obtener el mejor resultado. Cabe resaltar que, la estrategia de cada jugador está pues dada por las jugadas que éste puede hacer durante un juego. Esto es, una estrategia adecuada o ganadora estaría demostrando un comportamiento racional. En el ámbito de la TJ, una estrategia sería, de acuerdo al planteamiento de Morgenstern (1955): “una sucesión de jugadas que brindan a cada uno de los jugadores las mejores posibilidades, independientemente de lo que vaya a hacer el otro. (pág. 353)

Así, la relación entre el comportamiento racional y la mejor estrategia, podríamos ubicarla en el marco del conjunto de jugadas elegibles o movimientos electivos, cuya particularidad constituyen el marco de decisión en el cual los agentes epistémicos basan

* diana.ruiz@unach.mx

** francisco.rivas@unach.mx

el desarrollo de sus estrategias. A decir de Morgenster (1955), las “reglas del juego determinan cuáles son las posibles jugadas para cada uno, en qué caso y cuándo interviene el azar y cuándo un juego debe considerarse terminado” (pág. 352). Una forma de plantear el tipo de juego a jugar, dependerá entonces, de la información, ya sea diseñándose de forma disponible o no disponible para y entre los agentes epistémicos. Esto es, más allá del resultado de la interacción, debe considerarse el desarrollo del juego mismo. Lo anterior para considerar observarlo desde el ámbito de la argumentación, que resulta precisamente en la intención de la presente comunicación.

Ahora bien, pese a que los juegos que presenta Morgenster (y Von Newman) tienen una organización que los diferencia en: i) esenciales, que ofrecen ventaja a los participantes que establecen coalición y, ii) no esenciales, que no ofrecen una ventaja a los jugadores en coalición. Serán los primeros en los que se centrarán las investigaciones que busquen identificar las leyes fundamentales en las que se reconstruyan las distintas distribuciones y asignaciones en los conjuntos de solución en cada partida. De esta manera, “dominar significa pues, tener una preferencia bien definida” (pág. 362) y, donde las reglas del juego son dadas bajo las condiciones que representan los tipos de relaciones e interacciones que se establecen con o sin la información disponible de la que se hablaba hace algunas líneas.

Si bien este modelo intenta representar las relaciones económicas, resulta relevante para nuestra hipótesis el centrar nuestra atención en la interpretación resultantes. Mientras que, la teorización de los juegos busca desarrollar un modelo matemático que permita diseñar una estrategia ganadora, que resuelva las expectativas del comportamiento racional de los agentes involucrados, dicha interpretación puede dirigirse al campo de la argumentación, en tanto que, como proceso racional (estratégico) es un modo de decidir en torno a las razones que pueden darse o pedirse de forma lícita en un contexto de juego específico.

La Semántica Formal Estratégica (SFE) que propone Jaako Hintikka no contiene un carácter ontológico, sino que es de carácter lógico-pragmático en su representación semiótica, correspondiente a la semántica veritativa funcional de la lógica clásica. Debiendo proponer los inicios en la SFE desde la lógica proposicional nos permitirá identificar el valor de juego a nivel proposicional, que opera en las relaciones o condiciones iniciales dictadas por los conectivos lógicos. Los conectivos lógicos describen el tipo de relación que guardan proposiciones que se conjuntan, disyuntan, implican e inclusive se bi-implican o resultan en su dimensión semántica, expresivamente equivalentes. La negación, por su parte, para el lenguaje de orden cero de la lógica proposicional, la consideración del cambio de la variable proposicional o la sección del enunciado o proposición a contender. Al respecto, Jaako Hintikka propuso la Teoría Semántica de Juegos o Game-Theory Semantics (GTS), mediante la cual se presenta un cálculo de valores de verdad, así como elementos cuantificaciones, capaces de formalizar los distintos niveles de los lenguajes en los que se puede interpretar un modelo de razonamiento estratégico. Las condiciones

de significatividad de las “jugadas” que cada agente realiza se establecen en una serie de reglas que determinan el tipo de decisión posible que cada agente puede establecer. Según esta argumentación aquellas reglas determinan la semántica de la teoría de juegos, puesto que están encaminadas a determinar las condiciones de verdad o falsedad de cada una de las “jugadas”.

En la GTS el modelo que guía la semántica de juegos es el de estrategia, concepto que controla la elección individual de alguno de los dos “jugadores”. Las elecciones dependen de la información y movimientos que el jugador vaya haciendo. El hecho de que la información pueda verificarse o falsearse, es lo que permitirá en un desarrollo argumentativo les sea dado un contenido semántico asociado a la noción de estrategia adecuada y las vuelve susceptibles de ser posteriormente revisadas y, porque no, formalizadas. El papel central de la estrategia se define al considerar la verdad de una jugada como la utilidad esperada, en la que tal generación de una estrategia ganadora sea, en el cierre de una actividad argumental para el verificador inicial, la demostración de un razonamiento adecuado.

Al respecto puede enriquecerse la discusión, trayendo a colación la profundidad de la propuesta de Hintikka (1982), en el que su perspectiva señala que, para lenguajes de primer orden la GTS, funciona de la misma manera que la definición semántica de la verdad de Tarski, para oraciones atómicas. En el caso de la definición de las condiciones de verdad para oraciones no, la GTS aplica el mismo elemento de recursividad que la definición tarskiana de verdad logró para la semántica. Empero, la incursión de ésta temática, más allá del modelo matemático de formalización que se propone para el cálculo (limitado pues la semántica GTS es incompleta y sólo opera en lenguajes de primer orden), es una alternativa en el sentido de una Lógica Independiente Amigable (IF), que nos permite comprender cada una de las jugadas como un intercambio de información capaz de ser verificado o falseado; se relaciona con la viabilidad para aproximar dicho modelo como una propuesta entre el juego y la práctica semántica de la argumentación en espacios emergentes.

Lo que vuelve atractiva la GTS para nuestra propuesta es el hecho de que podemos contar con un criterio formal para determinar la validez de ciertas jugadas, con base en las condiciones de verdad de la información, lo que resulta muy útil a la hora de descartar ciertos movimientos o informaciones como inválidas o de nulo valor informativo.

Por último, lo que hemos denominado argumentación en espacios emergentes, no busca ubicarse en el ámbito de la retórica, sino dirigirse al análisis de la interacción y sus condiciones, por lo que todo espacio, dado o en construcción, en el que se encuentren al menos dos agentes epistémicos (considerando o no a la naturaleza como uno de ellos) podremos considerarlo un encuentro en el que se pone en juego, la defensa o ataque, la refutación o la verificación de una tesis. No partiendo de una razón práctica, sino estratégica, pero en cuyo carácter deliberativo se modele la argumentación y contra-argumentación, el dar y pedir razones, con el objeto entonces de elucidar el proceso, estudiar su dinámica y proponer formas más adecuadas de argumentar.

Referencias

- [1] Acero, J. J. (1986-1987). *La teoría de los juegos semánticos: dos enfoques*. *Theoria*, 427-459.
- [2] Acero, J. J. (1990). *Juegos en semántica: juegos semánticos*. *Enrahonar*, 65-85.
- [3] Hintikka, J. (1982). Game-Theoretical Semantics: Insights and Prospects. *Notre Dame Journal of Form Logic*, 219-241.
- [4] Hintikka, J. (2000). Game-Theoretical Semantics as a challenge to proof theory. *Nordic Journal of Philosophical Logic*, 4(2), 127-141.
- [5] Morgenstern, O. (1955). La teoría de los juegos y el comportamiento económico. *Económica*, 1(3-4), 345-375. Recuperado el 2021 de Marzo de 15, de <http://sedici.unlp.edu.ar/handle/10915/8865>
- [6] Perelman, Ch. (1997). *El imperio retórico. Retórica y argumentación*. Bogotá, Colombia: Norma.

LÓGICA, HISTORIA Y
FILOSOFÍA DE LA LÓGICA

El primer Tarski y la concepción universalista de la lógica

Víctor Aranda*
Universidad de Salamanca

Resumen

El objetivo de esta comunicación es cuestionar la tesis de que la cuantificación sobre un universo que es todo lo que existe y es, además, fijo (señalada como una de las características principales de la *concepción universalista* de la lógica) es incompatible con la metalógica. Para ello, mostraré que los teoremas demostrados por el primer Tarski que involucran las nociones de categoricidad y completud se discuten en el marco de una cuantificación fija que, por tanto, no está restringida a estructuras particulares. De hecho, este resultado es coherente con la tesis (que también suscribo) de que el concepto de *verdad en una estructura* no es explícitamente formulado por el primer Tarski.

A partir de unos seminarios que impartió Dreben en Harvard en los años 60, se popularizó la idea de que Frege, Whitehead y Russell tenían una *concepción universalista* de la lógica. Dicha concepción explicaría sus métodos y actitudes hacia la lógica y goza de mucha aceptación entre los historiadores de la disciplina. En particular, sus defensores argumentan que esta manera de concebir la lógica es incompatible con la formulación de cuestiones *metalógicas*.

Uno de los aspectos más característicos de la lógica “universal” es que el universo de cuantificación no es el dominio de una estructura concreta, sino que es todo lo que existe y es, además, fijo:

For Frege it cannot be a question of changing universes [...] His universe is the universe. Not necessarily the physical universe, of course, because for Frege some objects are not physical. Frege’s universe consists of all that there is, and it is fixed (Van Heijenoort, 1967, p. 325).

* vic.aranda@usal.es

In *Principia Mathematica* [...] within one type there is no restriction to a specific domain, and in that sense the universality is preserved. We have a stratified universe, but here again it is the universe, not a universe of discourse changeable at will (Van Heijenoort, 1967, p. 326).

Este rasgo de la concepción universalista de la lógica aparece repetidamente en numerosos comentarios a la obra de Frege y Russell, como en Goldfarb (1979) y en Hintikka (1988):

It is certainly true that Frege and Russell saw the quantifier as a central item in their logical systems [...] The ranges of the quantifiers –as we would say– are fixed in advance once and for all. The universe of discourse is always the universe, appropriately striated (Goldfarb, 1979, p. 351).

Y también en los trabajos de de Rouilhan (1991) y Rivenc (1993):

Cette logique est universelle au sens où elle s’applique à tout domaine d’objets pensables [...] le seul “univers du discours” est l’Univers tout court, le Monde, la totalité de l’être (Rivenc, 1993, p. 7).

Sin embargo, el universo de cuantificación de Tarski y Lindenbaum (1935) es todo lo que existe y es, además, fijo. La lógica que adoptan es la de los *Principia Mathematica*, modificada como sigue: se asume la teoría simple de tipos y el axioma de extensionalidad, se eliminan los operadores definidos y solo son sentencias las funciones proposicionales sin variables libres. El sistema contiene los axiomas de infinitud y elección.

Tarski (1934, p. 297) distinguía entre constantes lógicas, variables y constantes extra-lógicas. Entre las sentencias, destacan las que son “lógicamente deducibles”, esto es, las que son deducibles a partir *de la lógica*. Así, introduce una noción de “consecuencia”:

Let X be any set of sentences and y any sentence of the given theory. We shall say that y is a *consequence of the set X* of sentences, or that y is *derivable from the set X* , if either y is logically provable or if there is a logically provable implication having a sentence belonging to X , or a conjunction of such sentences, as its antecedent, and having y as its consequent (Tarski, 1934, p. 298).

Además, la lógica de Tarski y Lindenbaum (1935) contiene una “expresión simbólica” de la forma

$$R \frac{x', y', z', \dots}{x'', y'', z'', \dots}$$

que, según Tarski (1934, p. 310), significa lo mismo que la conjunción

$$R \in 1 \longrightarrow 1.V \sim_R V.x' \sim_R x''.y' \sim_R y''.z' \sim_R z'', \dots$$

donde V es la clase de todos los individuos, y $1 \longrightarrow 1$ es la clase de todas las relaciones uno-a-uno. R siempre denota una relación binaria entre x' y x'' . Por tanto, R será una relación uno-a-uno desde la clase de todos los individuos hacia sí misma ($V \sim_R V$) que “mapea” $x', y', z' \dots$ en $x'', y'', z'' \dots$, respectivamente. Es evidente que V no es el universo de una estructura particular, sino el universo *en general*.

Por último, una función proposicional es una expresión simbólica de la forma

$$\sigma(a, b, c, \dots; x, y, z, \dots),$$

donde a, b, c, \dots son constantes extra-lógicas y x, y, z, \dots son variables libres. La conjunción de los axiomas de una teoría será $\alpha(a, b, c, \dots)$.

El primer teorema que introducen Tarski y Lindenbaum (1935, p. 385) es el siguiente. Toda sentencia de la forma

$$(x', x'', y', y'', z', z'', \dots, R) : .R \frac{x', y', z', \dots}{x'', y'', z'', \dots} \cdot \supset : \sigma(x', y', z', \dots) \cdot \equiv \cdot \sigma(x'', y'', z'', \dots)$$

es lógicamente deducible.

Este teorema establece que para todo $x', x'', y', y'', z', z'', \dots, R$, si $x' \sim_R x'', y' \sim_R y'', z' \sim_R z'', \dots$, entonces las sentencias $\sigma(x', y', z', \dots)$ y $\sigma(x'', y'', z'', \dots)$ son lógicamente equivalentes. Esto es, si tenemos una relación uno-a-uno, R , entre x', y', z', \dots y x'', y'', z'', \dots , x', y', z', \dots satisface σ syss x'', y'', z'', \dots satisface σ . Adviértase que R no es un isomorfismo, sino un *automorfismo* (esto es, un isomorfismo de la clase de individuos en sí misma). En palabras de Tarski y Lindenbaum:

Th. 1 states that every relation between objects (individuals, classes, relations, etc.) [...] is invariant with respect to every one-one mapping of the “world” (i.e. the class of all individuals) onto itself and this invariance is logically provable (Tarski y Lindenbaum, 1935, p. 385).

Llegados a este punto, Tarski y Lindenbaum (1935, p. 390) define, entre otras, la noción de categoricidad para un sistema de axiomas $\alpha(a, b, c, \dots)$. Así, un sistema de axiomas $\alpha(a, b, c, \dots)$ es categórico syss la sentencia

$$(x', x'', y', y'', z', z'', \dots) : \alpha(x', y', z', \dots) \cdot \alpha(x'', y'', z'', \dots) \cdot \supset \cdot (\exists R) R \frac{x', y', z', \dots}{x'', y'', z'', \dots}$$

es lógicamente deducible.

En Tarski (1934), la definición de categoricidad es la misma, aunque hace especial hincapié en que es, en cierto sentido, una definición “fuerte” de categoricidad. Pues, en efecto, lo que están diciendo Tarski y Lindenbaum (1935) es que el sistema de axiomas $\alpha(a, b, c, \dots)$ es categórico syss, para todo $x', x'', y', y'', z', z'', \dots$, si x', y', z', \dots y x'', y'', z'' satisfacen α , existe un automorfismo entre x', y', z', \dots y x'', y'', z'' . Como señala Mancosu (2010), en teoría de modelos usamos un concepto de isomorfismo entre estructuras:

This is a rather strong condition and it is not quite the same notion of categoricity as used in contemporary model theory (where we have non-fixed domains and thus appealing to an automorphism of ‘the’ domain of individuals makes little sense) (Mancosu, 2010, p. 483).

Por tanto, una caracterización “fija” del universo de cuantificación no es incompatible con el desarrollo de cuestiones metalógicas.

Referencias

- de Rouilhan, P. (1991). De l’universalité de la logique. *L’âge de la science*, 93–119.
- Goldfarb, W. (1979). Logic in the twenties: the nature of the quantifier 1. *The Journal of Symbolic Logic*, 44(3), 351–368.
- Hintikka, J. (1988). On the development of the model-theoretic viewpoint in logical theory. *Synthese*, 77, 1–36.
- Mancosu, P. (2010). *The Adventure of Reason. Interplay between mathematical logic and philosophy of mathematics: 1900–1940*. Oxford: Oxford University Press.
- Rivenc, F. (1993). *Recherches sur l’universalisme logique. russell et carnap*. Payot, Paris.
- Tarski, A. (1934). Some methodological investigations on the definability of concepts. En *Logic, semantics, metamathematics: papers from 1923 to 1938* (pp. 296–319). Hackett Publishing.
- Tarski, A., y Lindenbaum, A. (1935). On the limitations of the means of expression of deductive theories. En *Logic, semantics, metamathematics: papers from 1923 to 1938* (p. 384–392). Hackett Publishing.
- Van Heijenoort, J. (1967). Logic as language and logic as calculus. *Synthese*, 17(1), 324–330.

An Analysis of Ancient Medical-Veterinary Diagnosis

Cristina Barés Gómez*
University of Seville

Hippiatry is the veterinary speciality which takes care of horses. We find its roots in the Ancient Near East. It is worth noting that not every text about horses are hippiatric texts. Indeed, a distinction must be made between horse's care texts (hippologic texts) and medical-veterinary texts (hippiatric texts)[6, 12], the difference of which is similar to the difference between physiology and diagnosis in medical science. As stressed by C. Bernard [4], both are important for the medical science, but they have not the same purposes. In this work, we analyse the structure of the hippiatric aspect of veterinary cares and the inferences underlying the diagnosis they describe. We claim that they are generally forms of abductive inference, triggered by an ignorance problem. That is, the starting point is a question triggered by the observation of a surprising fact (signs/symptoms), whose answer would be an illness, the hypothesis of which serves as a basis for further action in view of curing the patient-horse.

The hippiatric texts we find in Ancient Near East are Akkadian and Ugaritic texts. These texts need to be analysed in the context of the culture, and not only by comparing them, as well as in connection with the medical texts. That is to say, not only we must compare the Hippiatry with the Hippology, but we must also compare them with medical diagnosis. In this context, this work more precisely aims at a first analysis of the inferential structure of hippiatric texts in the light of the inferences used in the Mesopotamian medical practice.

This constitutes the first scientific medical-veterinary practice in History.

Although there is not, properly speaking, a medical practice in Ugarit, a wide amount of sources related to Akkadian medicine is available. Interesting is the fact that these Ancient medical texts seem to make use of the same kind of reasoning at stake in some modern medical diagnosis, namely abduction. Here, we understand abduction in terms of ignorance-preserving reasoning [8, 14, 11].

It is neither deduction, nor induction. The hypothesis of an illness is conjectured by the physician, who ends up acting despite the lack of definitive confirmation; i.e. by means of a full abduction [1, 3]. In Akkadian medicine, diagnosis is usually defined according to four parts: sign/symptoms, illness, prognosis and treatment. We also have, in some way, the same parts in the hippiatric texts we will analyse [9, 13, 1].

* crisbares@gmail.com

After having explained the Akkadian medical diagnosis in terms of abductive inference, we will offer an analysis of the Akkadian Hippiatry. We will use the BAM 159[10] text which, besides a medical diagnosis for persons, contains parts with illness treatment for horses. Both diagnosis (medical and veterinary) are located in the same texts in Akkadian literature which indicates they share the same goals as the roots of the medical science in Ancient Near East. On one side we have the medical science, and in the other we have the veterinary science.

This texts show us how the illness and its diagnosis seems to be quite known in persons, and maybe also in horses. The sign/symptoms might be similar, so recognizing the illness could be similar between persons and horses. The main difference would be the treatment. These texts confirm how the Akkadian Hippiatry is considered inside medical practice and in this regard as scientific practice, or better say, the origin of practical veterinary medicine. The difference at the inference level is that in the former (Akkadian medicine) we have a full abduction, while in the later (Hippiatry) we have just the last part of the inference (the treatment). As the diagnosis was similar, it is possible that hippiatric texts are only a part of the medical reasoning, the practical part of applying the specific treatment for horses.

As for the Ugaritic hippiatric texts, this structure results even more fuzzy.

We only have four texts (RS 5.285, RS 5.300, RS 17.120 y RS 23.484) [7, 5] and they have been classified inside “Scientific Texts” by the Semitic languages researchers [5]. Among the four texts, just one of them is practically complete, the RS 17.120 (KTU 1.85), so we will analyse this one. Here, we see that the illness and its name is completely absent, which does not happen in the Akkadian ones. The only aspect we have is the last part of the structure, the sign/symptoms and the treatment. We will analyse this structure by taking into account the contexts and three aspects: the linguistic conditional and the subjunctive inferential relation; the not so clear difference between the sign/symptoms and the illness; and the clear defeasibility of the relation between sign/symptoms and treatment.

After a formal analysis of the central structure of abduction in hippiatric Ugaritic diagnosis, we might see how the “explanatory” part of the abductive inference is completely absent. As the “explanatory” aspect of the scientific reasoning and concretely abductive reasoning has been considered as a fundamental part, we might consider that this diagnosis is not an abduction. Nevertheless, our point is precisely to mention that this “explanatory” aspect is not necessary, either sufficient in abduction. There is another part of the schema, we need to take into account as an important part, even when we deal with an “explanatory” abduction. This is exactly the part we have in Ugaritic and Akkadian Hippiatry, the further action which led us to a practical reasoning. We claim that they are perfectly abductive inference, although presented in somewhat different manners. At least, they explicitly manifest features of the last part of the diagnosis. We think they were correctly clasied as “Ugaritic Science”. In fact, they might be considered as the roots of the medical-veterinary diagnosis in Ancient Near East, although not with a full abduction

complete, with abductive inferential aspects, concretely with the activation of the hypothesis which correspond with the treatment [2].

References

- [1] C. Barés Gómez. Abduction in akkadian medical diagnosis. *Journal of Applied Logics - IfCoLog Journal of Logics and their Applications*, 5(8):1697-1722, 2018.
- [2] C. Barés Gómez. Lógica, Conocimiento y Abducción, chapter Un análisis de la inferencia en la práctica médico-veterinaria antigua. *Los textos hipiátricos de Ugarit.*, pages 265-284. College Publications, 2021.
- [3] C. Barés Gómez and M. Fontaine. Medical reasoning in public health emergencies: Below high standards of accuracy. *Teorema*, 40(1), 2021.
- [4] C. Bernard. *Introduction à l'étude de la médecine expérimentale*. Garnier-Flammarion, Paris, 1966.
- [5] P. Bordreuil and D. Pardee. *Manuel d'Ougaritique*, volume 2. Paul Geuthner, Paris, 2004.
- [6] C. Cohen. The ugaritic hipiatric texts and bam 159. *Journal of the Ancient Near Easter Society* 15, pages 1-12, 1983.
- [7] C. Cohen and D. Sivan. The Ugaritic Hipiatric Texts. A Critical Edition. *American Oriental Series, Essay* 9, 1983.
- [8] D. Gabbay and J. Woods. *The Reach of Abduction. Insight and Trial*. Elsevier, Amsterdam, 2005.
- [9] M. J. Geller. *Ancient Babylonian Medicine. Theory and Practice*. Wiley-Blackwell, Sussex, UK, 2010.
- [10] F. Köcher. Die Babylonisch - assyrische Medizin in Texten und Untersuchungen. De Gruyter, 1963, 80.
- [11] L. Magnani. Governing ignorance through abduction. *Logic Journal of the IGPL*, 2019.
- [12] D. Pardee. Trente ans de recherches sur les textes et les soins hipiatricues en langue ougaritique. Pallas 101. *Trousse du veterinaire dans l antiquite et au moyen age - instruments et pratiques*, 2016.
- [13] J. A. Scurlock and B. R. Andersen. *Diagnosis in Assyrian and Babylonian Medicine*. University of Illinois, 2005.
- [14] J. Woods. *Errors of Reasoning. Naturalizing the Logic of Inference*. College Publications, London, 2013.

Methodological Analysis for the Reconstruction of Language Families. The case of Semitic Languages

Cristina Barés Gómez*
University of Seville

Angel Nepomuceno-Fernández**
University of Seville

Francisco J. Salguero-Lamillar***
University of Seville

We usually analyze the reasoning and methodology of sciences by focusing in the kind of inference at stake. Induction and deduction are the most known logical inferences. Nevertheless, there is another kind of inference widely used in Science that use the hypothesis as its main aim. This type of inference is called abduction since Peirce [3]. Indeed, abduction is described by Peirce and other philosophers as a systematic hypothesis-forming operation. The basic and general aspect of abductive reasoning is that we have a background theory, at the starting point, there after a surprising fact (or several surprising facts, whose characteristics should be specified) appears which does not match with the background theory, so we need a hypothesis to explain this surprising fact/s. This schema presents many philosophical and logical problems which need to be solved. Nevertheless, we are not going into details about them, even if sometimes we will mention them. What we try to do in this paper is to analyze one particular kind of abduction in a concrete scientific research field, the reconstruction of a linguistic family group.

Deduction proves that something must be; Induction shows that something actually is operative. Abduction merely suggests that something may be (Peirce, CP 5.171) [3]. In fact, this kind of reasoning is one of the most used in science because it implies a conjecture and introduces new ideas. Nevertheless, the abductive result is something provisional. It can be questioned or denied. It is a conjectural solution, a provisional solution that makes us continue. The form of this type of inferences, in Peirce words, is the following:

* crisbares@gmail.com

** nepomu@us.es

*** salguero@us.es

The surprising fact, C, is observed;
But if A were true, C would be a matter of course,
Hence, there is reason to suspect that A is true. (Peirce, CP 5.189) [3]

Several problems remain with this notion of abduction. One of the main problems is: can or should we formalize abduction? If so, how? This work precisely tries to clarify some aspects of a particular type of abductive process by means of using certain formal tools. This point does not mean that the formalization is closed, neither unproblematic. We just try to get benefit from formalization to analyze some important aspects of abduction.

Our example is a reconstruction of a language family tree from their common Proto-Language. A Proto-Language is a hypothetical linguistic construction that would explain the evolution of a linguistic family. In fact, we conjecture the existence of an initial common language to be able to explain the difference among the languages that pertain to a family and the way they have evolved. Historical and comparative linguistics reconstructs forms, sounds, words and even sentence structures of a lost language by regarding the other language in the family tree. The point is that the reconstruction of a Proto-Language, or a complete family tree, could change because of new discoveries. Besides this, the conjecture of the existence of Proto-Languages has a further implication: we hypothesize the existence of a society that spoke those languages. The Proto-Language that we are going to use as an example for our methodological reconstruction of languages is Proto-Semitic [2].

In the second millennium, we have some problems establishing a clear branch of language history. In this case, we must follow the three facts that help to make an alternative hypothesis about the structure of this language family. First, we have the Amorite names that does not fit with the division in three branches. This evidence, as well as other linguistic phenomena, leads researchers such as von Soden [7] to the conclusion that in the second half of the third millennium B.C. there must have existed an older Semitic language. The name proposed by von Soden to designate this language is “Old-Amorite”. Second, we have a new unexpected fact, the discovery in 1929 of a new language. This language is similar to Hebrew and it was placed in the North-West Semitic branch. Nevertheless, it includes many archaic rests and similarities with Akkadian and Amorite. Hence, the question is whether Amorite and Ugaritic should be placed somewhere between the western and the eastern languages or belonging to the western branch. Third, we have in 1964 another discovery, the city of Ebla. Archaeologists have found there some texts written in Eblaite. This language could be dated around III millennium B.C. The point is that these three unexpected facts changed the hypothetical structure of the linguistic family theory about Semitic languages. So, a new hypothesis might emerge. That is: a common origin might have existed between North-West Semitic and East-Semitic that is already called by semiticians Paleo-Canaanite or Paleo-Syrian. This would be another hypothesis reached. The inference made for this kind of reasoning is the one we are going to analyze.

In order to give a clear account of our example, we should analyze it and clarify if we are in front of an analogical abduction [6] or a preduction [4]. Do we infer our hypothesis from the data obtained in other sciences to make a preduction? Or, do we make an analogy by means of inferences taken from other sciences and then we make an abduction? If we consider our example as analogical abduction, we have a relation (analogy) between settlements and languages. Nevertheless, this relation is based also on the concept of language that need to be spoken by persons. So, is it really an analogy between two facts or is it a deductive consequence of the concept of language? In order to highlight the problem, to establish if we are in front of an analogical reasoning, a preduction process, or another kind of selection in our example, we will use the mapping system for analogical reasoning from the multi-constrain theory by Holyoak and Thagard [1]. In addition, we are going to implement it in a modal framework in order to be able to account for inter-contextual selection of the hypothesis [5].

References

- [1] K. J. Holyoak and P. Thagard. The analogical mind. *American Psychologist*, 52(1):34-44, 1997.
- [2] E. Lipinski. *Semitic Languages Outline of a Comparative Grammar*. *Orientalia Lovaniensia Analecta*, Leuven-Paris-Sterling (Virginia), second edition, 2001.
- [3] C.S. Peirce. *Collected Papers of Charles Sanders Peirce*. Harvard University Press, Cambridge, 1931-1958.
- [4] A. Rivadulla. *Complementary strategies in scientific discovery: Abduction and Prediction*. *Nordic Studies in Pragmatism*, 2010.
- [5] F. Soler-Toscano, D. Fernández, and A. Nepomuceno-Fernández. A modal framework for modelling abductive reasoning. *Logic Journal of the IGPL*, 20(2):438-444, 2010.
- [6] P. Thagard. *Computational Philosophy of Science*. Massachusetts Institute of Technology, 1993.
- [7] W. von Soden. Zur einteilung der semitischen sprachen. *Wiener Zeitschrift für die Kunde des Morgenlandes*, 56:177-191, 1960.

Essentialist Negative Kinds and Internal Negations

Javier Belastegui*
University of the Basque Country

Abstract

Some philosophers (among them David Armstrong), have argued that there are no negative natural properties, that the collection of natural properties is not closed under negation. If there are no negative natural properties, then plausibly there are no negative kinds either. The aim of this talk is to argue that, under essentialist assumptions concerning the membership and identity conditions of kinds, this is not the case. There must exist negative essentialist kinds. Moreover, the behavior of these negative kinds is non-classical, for instance, an object may fail to belong both to a kind K and to its negative kind $\neg K$. Thus these negations correspond to internal or term negations, not to propositional ones. In order to argue for my claim I will extend an algebraic model for kinds originally introduced by Richmond Thomason by using Rudolf Wille's Theory of Concept Lattices. This model will be shown to preserve the essentialist membership and identity conditions of kinds and to be committed to the existence of negative kinds.

Keywords: Natural Kinds, Natural Kind Essentialism, Natural Properties, Negative Properties, Formal Models of Properties and Kinds.

1 Negative Kinds and Internal Negations

Some philosophers, including famously David Armstrong [1], have suggested that natural properties are not closed under Boolean negations. If P is a natural property, then its negation $NOT(P)$ should have as an extension the complement of the extension of P . However, if P is a natural property, $NOT(P)$ need not be a natural property too. For example, although *BEING RED* or *HAVING MASS R* are natural properties, *NOT BEING RED* and *NOT HAVING MASS R* are not. Natural or sparse properties are the basis for resemblance and causal relations among objects and they are supposed to play a role either in measurement or in natural laws. Not red things are not sufficiently similar to

* fjbelastegui@gmail.com

each other, not being red provides no causal powers to not-red things and not-redness is not appealed to in order to give any sort of theoretical explanation.

But if natural properties are not closed under negation, this might suggest that natural *kinds* are not closed under negation either. Although *BIRD* or *ELECTRON* are kinds, *NOT BIRD* and *NOT ELECTRON* need not be kinds. Not bird things are not sufficiently similar to each other, not birdness provides no causal powers to not-bird things and not-birdness is not appealed to in order to give any sort of serious theoretical explanation. In particular, not-birdness (or not being an electron) is not mentioned by classifications, either in science or ordinary talk.

The aim of this talk is to argue that, under an essentialist conception of natural kinds this cannot be the case: under minimal assumptions, the essentialist is forced to accept the existence of *negative kinds*. Moreover, these negative kinds do not behave classically. For instance, they do not satisfy the Excluded Middle.

It is known that Aristotle distinguished between two kinds of negation, namely *external* or *propositional negation* and *internal* or *term negation*. To use a classical example, the proposition that *a* is not mortal involves a propositional negation, it says that the term ‘mortal’ does not apply to the object *a*. The proposition that *a* is non-mortal (immortal) involves a term negation, for it says that the negative term ‘immortal’ applies to the object *a*. If *a* is non-mortal (or immortal), then it follows that *a* is not mortal. However, the converse does not hold, since there are entities to which neither the predicate ‘immortal’ nor the predicate ‘mortal’ apply (e.g. stones or machines). Contrary to propositional negation, internal negation is not classical, because it does not satisfy the Excluded Middle. This distinction was also present in Kant’s classification of logical judgements according to quality into affirmative, negative and infinite [6] (e.g. see [9]). Whereas negative judgements give rise to propositional negations, infinite judgements correspond to internal negations.

2 Negative Essentialist Kinds

But what does all this have to do with kinds? It turns out that the essentialist is committed precisely to the sort of negative kinds that Aristotle and Kant were describing. There are many different essentialisms of kinds such as those by Ellis [4], Lowe [7], Bird [2] or Tahko [10]. Nevertheless, all essentialists will agree at least on two claims. One, that all the instances of a kind share a *general essence*. Sharing a certain general essence with other objects will force all the instances of a kind to exemplify certain natural attributes essentially. If one is to give a *real definition* of a kind one will have to appeal, sooner or later, to those attributes essentially exemplified by all the instances of the kind. These

attributes are taken to be necessary and sufficient conditions for membership to the kind. So these attributes fix the *membership conditions* of the natural kinds. Second, that under essentialist assumptions, *identity conditions* for kinds themselves should be fixed by their general essences. The claims can be put as follows:

Essentialist Membership Conditions: For each natural kind K , there are some natural attributes P_1, \dots, P_n which are such that, for any object x , x instantiates K iff x essentially exemplifies P_1, \dots, P_n .

Identity Conditions for Kinds: Two kinds K, K' are identical iff they have the same general essence.

The sort of essentialism I will consider defends also the *converse* of the first claim, namely that if an entity K is such that it satisfies the previous conditions, then K is a natural kind. In other words, the two directions taken together give an essentialist definition of natural kinds.

To argue for the claim that there must be essentialist negative kinds, I will introduce a formal model for kinds by making use of Wille's Theory of Concept Lattices [5]. Lattices are algebraic structures that have already been used to model kinds, first by Thomason [11] and also as part of the semantics for syllogistic logic by Corcoran [3] and Martin [8]. Although these models give some information about how kinds are related by specificity relations, they are silent regarding how the membership conditions of kinds are fixed by the natural properties shared by their instances. In contrast, the kind of lattices I will make use of provide an analysis of a kind into two components. Given a set of objects S , a set of natural attributes Q and a relation of essential exemplification $I \subseteq S \times Q$, a pair $A \subseteq S$ and $B \subseteq Q$ forms a *natural kind* $K = (A, B)$ iff:

$$A = \{x \in S \mid \forall P \in B \ xIP\} \qquad B = \{P \in Q \mid \forall x \in A \ xIP\} \qquad (1)$$

Thus an essentialist kind is represented as a pair of sets (A, B) which is such that A is the *extension* of B , the set of all the objects that essentially exemplify all the natural attributes in B , whereas B is the *intension* of A , the set of all the natural attributes that are essentially exemplified by all the objects in A . Some metaphysicists will balk at this set-theoretic modelling of kinds. Don't we already know that properties (or kinds) and sets have very different existence and identity conditions? Fair enough, but the point of a good model is to have enough power to represent adequately those features of the phenomenon under investigation, in this case, essentialist natural kinds. In the case that concerns us, note that these set-theoretic entities preserve precisely the two principles we introduced

before. First, kinds are constituted by certain objects essentially exemplifying certain natural attributes, in such a way that the latter fact fixes the membership conditions of the kind. Second, two kinds are identical iff they have the same intension iff they have the same general essence.

The set of all the natural kinds forms a complete lattice (known as a ‘concept lattice’ [5]). By using a definition by Wille [12], it can be shown that for each kind K there are two candidates for negative kinds namely:

$$\neg K = (e(B^c), ie(B^c)) \qquad \star K = (ei(A^c), i(A^c)) \qquad (2)$$

We call $\neg K$ the *opposition* of K and $\star K$ the *negation* of K . Whereas $\neg K$ contains those objects that essentially exemplify all the attributes that are not in the intension of K , the intension of $\star K$ contains all those attributes that are not essentially exemplified by all the objects in K . In other words, the general essence of $\neg K$ includes all the attributes that are not in the essence of K and possibly more, whereas the general essence of $\star K$ includes all those attributes that are essentially shared by objects that are not K -s. It will be shown that, whereas \neg does not satisfy the Excluded Middle, \star does not satisfy Non-Contradiction (with respect to the ‘disjunction’ and ‘conjunction’ operations algebraically defined in the lattice). Thus, the behaviour of these negations is non-classical.

It is worth noting that the lattice of kinds is closed under these two operations. The existence of negative kinds follows from the very definition of natural kinds as extension-intension pairs. Thus the essentialist is forced to accept not one but two sorts of negative kinds. A fortiori, the suggestion that there are no negative natural kinds is false, at least under an essentialist conception of natural kinds.

References

- [1] Armstrong, D. M. (1989): *Universals: An Opinionated Introduction*. Boulder: Westview Press.
- [2] Bird, A. (2015): The Metaphysics of Natural Kinds. *Synthese*, 1–30.
- [3] Corcoran, J. (1972): Completeness of an Ancient Logic. *The Journal of Symbolic Logic*, **37**, 4, 696–702.
- [4] Ellis, B. (2001): *Scientific Essentialism*. Cambridge: Cambridge University Press.
- [5] Ganter, B.; Wille, R. (1999): *Formal Concept Analysis: Mathematical Foundations*. Springer-Verlag: Berlin and Heidelberg.

- [6] Kant, I. (1988; 1800): *Logic*. Translated by R. S. Hartmann; W. Schwarz from *Immanuel Kants Logik: Ein Handbuch zu Vorlesungen*. New York: Dover Publications.
- [7] Lowe, J. (2006): *The Four-Category Ontology*. Oxford: Oxford University Press.
- [8] Martin, J. N. (1997): Aristotle's Natural Deduction Reconsidered. *History and Philosophy of Logic*, **18**, 1, 1–15.
- [9] McLaughlin, P.; Schlaudt, O. (2020): Kant's Antinomies of Pure Reason and the 'Hexagon of Predicate Negation'. *Logical Universalis*, **14**, 51–67.
- [10] Tahko, T. (2015): Natural Kind Essentialism Revisited. *Mind*, **124**, 495.
- [11] Thomason, R. (1969): Species, Determinates and Natural Kinds. *Nous*, **3**, 1, 95–101.
- [12] Wille, R. (2000): Boolean Concept Logic. In B. Ganter, G. W. Mineau (eds.) *International Conference on Conceptual Structures 2000*. Berlin, Heidelberg: Springer-Verlag.

A solution to the problem of mixed inferences by juxtaposition

Carlos Benito*
University of Barcelona/Logos

José Martínez**
University of Barcelona/Logos

The discussion on logical pluralism has been thriving within philosophy of logic in recent years. Probably, this great interest is due to the importance of some issues that are involved in the debate around pluralism, such as the normativity of logic, its formality or even its chief aim as a discipline.

Logical pluralism can be characterized as the thesis that there are at least two legitimate and extensionally non-equivalent theories of logical validity, i. e. of which conclusions follow logically from which premises.

Now, there are alternative strategies for motivating and arguing for logical pluralism, some of those giving rise to different kinds of pluralism. The most popular of those kinds is the one articulated by JC Beall and G. Restall (2006), whose essential aspect is the generalization of the Tarskian account of validity, namely:

Generalised Tarski Thesis (GTT): An argument is valid_x if and only if, in every case_x in which the premises are true, so is the conclusion.

Their pluralist thesis consists in arguing that there are various equally legitimate specifications of case_x, which give rise to different notions of validity. Therefore, they claim that there are at least two correct logics, which are all-purpose¹, normative, necessary and formal.

The version of logical pluralism that is going to concern us here is a different one, though. We will focus on logical localism and, especially, on the most important challenge that has been raised against it, which is the problem of mixed inferences. The localist thesis states that there is a multiplicity (or at least two) of domains of discourse, in such a way that correct reasoning in these domains takes the form of different logics. In other terms, that there are various domains that ask for different logical theories to capture correct reasoning within those domains.

* carlos.b.21@hotmail.com

** jose.martifer@gmail.com

¹ In the sense of Field (2009), i. e. logics that apply to any domain of reasoning.

This thesis can be motivated and supported on different grounds, but a particularly illuminating one is the argument that concludes in logical localism by defending alethic pluralism, this is, by defending that there is more than one way of being true depending on the objects that our discourse is about and the types of its concepts. Thus, alethic pluralism aims at providing an explanation to the intuition that “there are two apples on the table”, “every even whole number greater than two is the sum of two prime numbers”, “it is morally wrong to have a patent for the vaccine against COVID-19” or “that black joke wasn’t funny” have different ways of being true.

The route from alethic pluralism to logical localism is natural and obvious: if there is more than one way of being true and logical consequence is the necessary preservation of truth in virtue of logical form, then it is natural to expect that different logics could be correct in different domains. For example, if correspondence theory is adequate for the physical domain and an epistemic notion of truth is the best explanation of the truth in ethics, this could plausibly entail that classical logic is correct in physics but intuitionistic logic is correct in ethics. We will accept that these logics are correct for those domains for the sake of the discussion.

This approach faces the problem of mixed inferences (Tappolet 1997). Consider “Either murder is wrong or snow isn’t white. Snow is white. Therefore, murder is wrong”. This argument is intuitively valid, even though premises and conclusion do not belong to the same domain. How should we then explain its validity?

Lynch (2008) proposes that, in cases of mixed inferences, one adopts a modest attitude: the valid rules are the ones that are valid in each of the domains of the sentences occurring in the argument. The criterion is justified because, if we use a stronger logic, we could end up accepting sentences that are false in their domain. Consider the following example presented by Lynch (2008): “If it is not the case that Sophie’s choice is morally right, then grass is not green. But grass is green; so Sophie’s choice is morally right”. Let us suppose that we do not accept as true either “Sophie’s choice is morally right” or its negation. The argument is classically valid and intuitionistically invalid. If we accepted the stronger logic, we would be forced to accept its conclusion, even though (by assumption) we want to consider the conclusion neither true nor false.

This modest criterion has been criticised by Wrenn (2018), because there are arguments that are intuitively valid and that Lynch’s criterion rejects as invalid. As an example, consider this argument: “Either murder is not wrong or it’s not the case that snow isn’t white. Murder is wrong. Therefore, snow is white.” (1) According to the modest criterion, the intersection of the logics of the domains is intuitionistic logic, and the argument is intuitionistically invalid. However, since disjunctive syllogism is valid in intuitionistic logic and double-negation elimination is valid only in classical logic, but it is applied here to a sentence of the physical domain, the argument should be valid after all.

We agree that this is a correct criticism of the modest criterion. Wrenn (2018) proposes possible improvements of the modest criterion and finds new intuitively valid arguments

that the modified criteria cannot consider valid. Hence, lacking a solution to the problem of mixed inferences, Wrenn concludes that, if one accepts alethic pluralism, one should reject logical localism. Actually, we contend that the problem of mixed inferences is just as damaging even if logical localism is motivated not by alethic pluralism, but in a different way².

Our aim in the talk is to offer a solution to the problem of mixed inferences that improves on the modest criterion and its variations. We think that the search for solutions has been hindered by the disregard of the formal tools developed by logicians working on the combination of logics. In this work we will use juxtaposition, a method for the combination of logics created by J. Schechter (2011). In this method, the logical connectives are also specific to domains. We think that this represents more faithfully the position of localism and it opens new questions: is the disjunction in the first premise of (1) classical or intuitionistic? How do the connectives from the different logics interact among them?

With the use of the sophisticated tools offered by juxtaposition, we will show that the logical form of the inference (1) is valid. We will also justify as valid other counterexamples presented by Wrenn (2018).

References

- [1] Beall, J. C. and G. Restall (2006). *Logical Pluralism*. Oxford: Clarendon Press.
- [2] Field, H. (2009). "Pluralism in Logic". *The Review of Symbolic Logic* 2: 342-59.
- [3] Lynch, M. P. (2008). "Alethic Pluralism, Logical Consequence and the Universality of Reason". *Midwest Studies in Philosophy* 32: 122-40.
- [4] Schechter, J. (2011). "Juxtaposition: a new way to combine logics". *The Review of Symbolic Logic* 4: 560-606.
- [5] Tappolet, C. (1997). "Mixed inferences: A problem for pluralism about truth predicates". *Analysis* 57: 209-10.
- [6] Wrenn, C. B. (2018). "A Plea for Immodesty: Alethic Pluralism, Logical Pluralism, and Mixed Inferences", in Wyatt, J., Pedersen, N. J. L. L. and Kellen, N., editors. *Pluralisms in Truth and Logic*. Palgrave Macmillan, 387-406.

² For instance, via Quinean-inspired considerations on fundamental inference rules accepted in different domains, for philosophers that disregard semantic explanations of logical consequence.

La construcción del espacio lógico de posibilidad

Luis Alberto Canela Morales*
Universidad Nacional Autónoma de México

Introducción y planteamiento del problema

Según lo dicho por Wittgenstein en el *Tractatus*, el mundo es la totalidad de los hechos, no de las cosas. Si por “hecho” entendemos una manera de ser del mundo tal que el mundo es de esa manera, bien podríamos decir que los enunciados que componen el corpus del conocimiento son, en realidad, esas posibles maneras de ser del mundo y, en cierta medida, de estas condiciones de posibilidad depende el conocimiento que del mundo se tenga. En efecto, el conocimiento de la totalidad de los hechos reclama una correlación coherente y consistente con su propio campo de posibilidad, esto es, con el conjunto de postulados que, efectivamente, poseemos como sujetos epistémicos. De lo que se trata es, pues, de comprender las múltiples maneras de ser del mundo en virtud de acontecimientos fácticos, y no en la consideración de la mera y pura posibilidad. En ese sentido, y contrario a lo que también señala Wittgenstein, “conocer un objeto significa conocer todas las posibilidades de su ocurrencia en estados de cosas” (2009, §2.0123), no todas las posibilidades son hechos.

Así pues, en la construcción del conocimiento se reconoce el papel que juega el lenguaje, las creencias y, por supuesto, su lógica. En este caso, de un tipo particular de lógica: la lógica modal junto con la metafísica de la teoría de los mundos posibles y su pertinencia para la comprensión de las proposiciones de la ciencia. El problema que quiero abordar, radica, precisamente, en el uso del operador de “posibilidad” encuadrado en la noción de mundos posibles y su relación con los enunciados epistémicos. De manera particular, mi interés es presentar una posible solución al problema de afirmar que toda posibilidad lógica, usada en los enunciados matemáticos, siempre debe estar en relación con las maneras efectivas de ser del mundo. Para lograr lo anterior, me permitiré discutir ciertos aspectos e implicaciones significativas de la propuesta del filósofo analítico Agustín Rayo (MIT) denominada la construcción del espacio lógico de posibilidad. El espacio de posibilidad, según Rayo, es el conjunto de alternativas con las que trabajamos cuando nos preguntamos cómo es el mundo. Dicho de otra manera, es el conjunto de distinciones entre las maneras en las que podría ser el mundo o sobre las maneras en las que de hecho es. El espacio lógico incluye todas esas maneras de ser del mundo, las maneras en las que de hecho es y las maneras en las que meramente podría ser. Ahora bien, parte de lo que se hace cuando construimos nuestra concepción del espacio de posibilidad tiene que ver con la aceptación de ciertos enunciados de identidad (“que p sea el caso es simplemente

* luisanela25@gmail.com

que q sea el caso”). Así, el enunciado de identidad “que haya agua es simplemente que haya H₂O” impone un límite a nuestro espacio de posibilidad porque nos dice que no hay tal cosa como una posibilidad en donde hay agua, pero no H₂O o viceversa. Un ejemplo particular: “que el número de los dinosaurios sea cero es simplemente que no haya dinosaurios”. Alguien que rechaza esta proposición es alguien que cree que del hecho de que no haya dinosaurios no se sigue que el número de los dinosaurios es cero. En cambio, alguien que acepta dicha proposición cree que el mero hecho de que no haya dinosaurios ya se hizo todo lo que se tenía que hacer para que el número de los dinosaurios sea cero.

La argumentación es la siguiente: de acuerdo con la epistemología de los enunciados de identidad propuesta por Rayo, cuando uno acepta un enunciado de identidad el resultado es que se elimina cierto espacio teórico. Efectivamente, quien pregunta por la justificación de lo anterior en realidad está introduciendo una presuposición falsa a la discusión, pues lo que hacemos es que concebimos el significado como incrustado en el mundo, este último permite que los objetos mundanos y abstractos aparezcan con el significado que tienen. Por esta razón, según Rayo, cada objeto debe comprenderse necesariamente en su correlación con la subjetividad que está implicada en una relación intencional con el campo de donación inmediata que es el mundo.

La ponencia se dividirá en dos partes. La primera describirá la noción de espacio lógico de posibilidad en la obra de A. Rayo y la segunda profundizará en la posible relación que esta tiene con la propuesta fenomenológica de Husserl.

Bibliografía principal

Tomos de Husserliana

[Hua XII] *Philosophie der Arithmetik. Mit ergänzenden Texten (1890-1901)*. Hrsg. Lothar Eley, The Hague, Martinus Nijhoff, 1970.

[Hua XXII] *Aufsätze und Rezensionen (1890-1910)*. Hrsg. Bernhard Rang, The Hague, Martinus Nijhoff, 1979.

Agustín Rayo

— (2002). “Frege’s unofficial arithmetic”. *The Journal of Symbolic Logic*, 67, 1623-38.

— (2003). “Success by default?” *Philosophia Mathematica*, 305-22.

— (2005). “Logicism reconsidered”.

— (2008). “On specifying truth-conditions”. *The Philosophical Review*, 117, 385-443.

— y Uzquiano, Gabriel (Eds.) (2006). *Absolute Generality*. Oxford University Press, Oxford.

— (2009a). “Vague Representation”, 117, pp. 329-373.

— (2009b). “Towards a Trivialist Account of Mathematics”, En Bueno y Linnebo (Eds.), *New Waves in Philosophy of Mathematics*, Palgrave McMillan.

- (2012). “Hierarchies Ontological and Ideological” (con Øystein Linnebo), *Mind* 121, pp. 269-308, 2012.
- (2013). *The Construction of Logical Space*. Oxford: Oxford University Press, 2013. (192 pp.)
- (2015). *La Construcción del Espacio de Posibilidades*. Instituto de Investigaciones Filosóficas, UNAM, 2015.
- (2019). *On the Brink of Paradox*. Cambridge, MIT Press, 2019.

Bibliografía general

- [1] Bégout, Bruce (2000). *La genéalogie de la logique. Husserl, l’antéprédicatif et le catégorial*. Paris: Vrin.
- [2] Benacerraf, Paul y Putnam, Hilary (Ed.) (1983). *Philosophy of Mathematics*. Cambridge University Press, Cambridge.
- [3] Benoist, Jocelyn (2002a). *Entre acte et sens. Recherches sur la théorie phénoménologique de la signification*. Paris: Vrin.
- [4] Burgess, John (1997). *A Subject with No Object: Strategies for Nominalistic Reconstruction of Mathematics*, Oxford University Press.
- [5] Carnap, Rudolf (1988). *La construcción lógica del mundo*. México, UNAM.
- [6] — (1998). *Filosofía y sintaxis lógica*, México, UNAM.
- [7] Centrone, Stefania (2010). *Logic and Philosophy of Mathematics in the Early Husserl*, Dordrecht: Springer.
- [8] — (Ed.) (2017). *Essays on Husserl’s Logic and Philosophy of Mathematics*. Dordrecht: Springer.
- [9] Cobb-Stevens, R. (1998). *Husserl et la philosophie analytique*. Paris: Vrin.
- [10] Courtine, Jean-François, ed. (1996). *Logique et phénoménologie*. Paris: Éditions Rue d’Ulm.
- [11] Da Silva, Jairo José (1993). “Husserl’s Philosophy of Mathematics”. *Manuscrito* 16.2, pp. 121–148.
- [12] — (1999). “Husserl’s Conception of Logic”. *Manuscrito* 22.2, pp. 367–397.
- [13] Field, Hartry (1980). *Science Without Numbers*. Basil Blackwell and Princeton University Press, Oxford and Princeton.
- [14] Frege, G. (1884). *Grundlagen der Arithmetik. Eine logisch mathematische Unter-suchung über den Begriff der Zahl*. Breslavia: Koebner.

- [15] — (1980). *Philosophical and Mathematical Correspondence*. Oxford: Basil Blackwell. (Trad.) Hans Kaal, Gottfried Gabriel et al., (Eds.)
- [16] Haaparanta, Leila, (Ed.) (1994). *Mind, Meaning and Mathematics. Essays on the Philosophical Views of Husserl and Frege*. Dordrecht, Boston, London: Springer.
- [17] — (Ed). (2009). *The Development of Modern Logic. Essays on the Philosophical Views of Husserl and Frege*. Oxford University Press: New York.
- [18] Hale, Bob and Wright, Crispin (2001). *The Reason's Proper Study: Essays towards a Neo-Fregean Philosophy of Mathematics*. Clarendon Press, Oxford.
- [19] Lewis, David (1986). *On the Plurality of Worlds*. Blackwell, Oxford and New York.
- [20] Lohmar, Dieter (1989). *Phänomenologie der Mathematik. Elemente einer phänomenologischen Aufklärung des mathematischen Erkenntnis nach Husserl*. Dordrecht: Kluwer.
- [21] Maddy, Penelope (1997). *Naturalism in Mathematics*. Oxford: Oxford University Press.
- [22] Mancosu, Paolo (Ed.) (1998). *From Brouwer to Hilbert. The Debate on the Foundations of Mathematics in the 1920's*. Oxford University Press: New York.
- [23] Miller, J. Philip (1982). *Numbers in Presence and Absence: A Study of Husserl's Philosophy of Mathematics*. The Hague /Boston/London: Martinus Nijhoff.
- [24] Mohanty, J.N. (1982). *Husserl and Frege*, Bloomington: Indiana University Press.
- [25] Morton, Adam y Stich, Stephen (Eds.) (1996). *Benacerraf and his Critics*. Basil Blackwell, Oxford.
- [26] Ortiz Hill, Claire (1991). *Word and Object in Husserl, Frege and Russell. The Roots of Twentieth-Century Philosophy*. Ohio: Ohio University Press.
- [27] — y Guillermo Rosado Haddock (2000). *Husserl or Frege? Meaning, Objectivity and Mathematics*. La Salle: Open Court.
- [28] — y Da Silva, Jairo José (2013). *The Road not Taken. On Husserl's Philosophy of Logic and Mathematics*. United Kingdom: College/Lighting Source/ Milton Keynes.
- [29] Parsons, Charles (1983). *Mathematics in Philosophy*. Cornell University Press, Ithaca, NY.
- [30] Salmon, Nathan U. (1995). *Reference and Essence*. Prometheus Books, Amherst, NY.
- [31] Stalnaker, Robert C. (1984). *Inquiry*. MIT Press, Cambridge, MA.
- [32] — (1999). *Context and Content*. Oxford University Press, Oxford.

- [33] Tieszen, Richard L. (1989). *Mathematical Intuition. Phenomenology and Mathematical Knowledge*. Dordrecht, Boston, London: Kluwer.
- [34] — (2005). *Phenomenology, Logic and the Philosophy of Mathematics*. Cambridge: Cambridge University Press.
- [35] — (1984). *Husserl and Realism in Logic and Mathematics*. Cambridge: Cambridge University Press.
- [36] van Atten, Mark (2007). *Brouwer meets Husserl. On the phenomenology of Choice Sequences*. Netherlands: Springer.
- [37] — (2015). *Essays on Gödel's Reception of Leibniz, Husserl and Brouwer*, Heidelberg/New York: Springer.
- [38] Wright, Crispin (1983). *Frege's Conception of Numbers as Objects*. Aberdeen University Press, Aberdeen.

Igualitarismo y Realismo en la metaontología de Hirsch

Pablo Carrete Varela*
Universidad de Valladolid

Introducción

En el marco de las recientes discusiones ontológicas y metaontológicas existe cierto grupo de teorías deflacionistas cuyo máximo representante es Eli Hirsch, referidas mediante el rótulo “*Quantifier Variance*”. Él es autor de la teoría metaontológica más refinada de este tipo (Hirsch 2002a). De hecho, es común ver usos del término “*Quantifier Variance*” asociándolo a la teoría metaontológica de Hirsch por su lado, como una de las tesis que componen el grueso de su teoría (la de la variación cuantificacional); incluso el propio Hirsch, usa la expresión en ese doble sentido tal y como indicaré. En este texto utilizaremos la expresión ‘variación cuantificacional’ para referirnos a la tesis, no para referirnos a la postura metaontológica. En cualquier caso, el objeto del análisis subsiguiente es mostrar que ciertos componentes de la propuesta de Hirsch son incompatibles entre sí, a saber: el igualitarismo, el realismo y el principio (**T-weak**).

La teoría

La teoría de Hirsch parte de dos presupuestos metaontológicos y de tres tesis. Los presupuestos son el pluralismo, al respecto de la cuantificación, y el igualitarismo de los lenguajes ontológicos¹ (Hirsch 2011, pp. xiii-xv). Las tesis (Hirsch 2009) son: *variación cuantificacional*, *disputas verbales* y *lenguajes alternativos*. El presupuesto igualitarista es relevante, pues es el origen del problema que expondré. Ésta es la definición² del igualitarismo de Hirsch (2011):

Quantifier variantism says that there is no uniquely best ontological language with which to describe the world. I take this to imply that (i) there are a number of possible truth-conditionally equivalent ontological languages, and (ii) these languages are of equal metaphysical merit (p. xiv).

* pablo.carrete96@gmail.com

¹ Por lenguajes ontológicos entenderé, simplemente, lenguajes empleados en el marco de una teoría ontológica.

² Aquí Hirsch emplea el primer uso de “variación cuantificacional”.

La idea es que existen numerosos lenguajes ontológicos posibles, los cuales son equivalentes en términos de condiciones de verdad, y tales lenguajes son, desde una perspectiva metafísica, “igualmente meritorios”. Cabe destacar dos cosas del fragmento: (a) la idea de que los lenguajes ontológicos *describen el mundo* y (b) clarificar en qué sentido son “igualmente meritorios”.

En lo tocante a (a), cabe percatarse de que la teoría de Hirsch se sustenta sobre un gran compromiso realista (Hirsch 2002a, p. 52). El filósofo se esfuerza en distanciarse de cualquier tipo de antirrealismo. A pesar de ello, Hirsch es vago al caracterizar su compromiso, se limita a expresar ese rechazo:

Consider the two sentences: ‘There exists something composed of Clinton’s nose and the Eiffel Tower’ and ‘Whether or not there exists something composed of Clinton’s nose and the Eiffel Tower depends on our linguistic decisions.’ Quantifier variance implies that the expression ‘there exists something’ can be interpreted in a way that makes the first sentence true or in a way that makes it false. But there is no relevant way to interpret ‘there exists something’ that would make the second sentence true. The second sentence expresses an absurd form of linguistic idealism that is not at all implied by quantifier variance (Hirsch 2002a, p. 52).

El fragmento expresa la tesis de la *variación cuantificacional*. Según la variación cuantificacional, el significado de la expresión “existe” puede variar entre diferentes lenguajes, donde tal término y su funcionamiento se identifican con el cuantificador existencial de la lógica cuantificacional *clásica* en todos ellos (Hirsch 2002a, pp. 53-54). Tomando la variación cuantificacional y sosteniendo el punto (i) de la primera cita es posible que cierto mereólogo afirme verdaderamente que “el *mesavaso*³ que tengo delante está lleno de agua”, lo que otros filósofos negarían, sosteniendo que lo adecuado es afirmar que “el vaso situado encima de la mesa que tengo delante está lleno de agua”. La primera de las oraciones es defendible por un mereólogo que acepta la tesis de la *composición irrestricta* (Lewis 1991, pp. 7-8). La segunda oración podría sostenerla un filósofo del lenguaje ordinario (en adelante *sensatistas*), como el propio Hirsch (2002b, pp. 104-105). El realismo de Hirsch implica, además, que las condiciones de verdad vienen determinadas por la referencia al mundo, tal como refleja la primera cita.

En cuanto a (b), es preciso aclarar que implica que lenguajes como los anteriores son igualmente meritorios. La forma más simple de entender la idea es que, en sus respectivas ontologías, las dos oraciones ejemplificadas son verdaderas (Hirsch 2002a, pp. 54-56). Siguiendo la cita de Hirsch, ambas descripciones de la realidad son igualmente adecuadas en términos de sus condiciones de verdad, tal y como indica (i). La noción de *lenguaje expresivamente empobrecido* de Eklund ayuda a entender esta idea: “A language is expressively impoverished just in case there are facts which the language does not possess the means

³ En términos mereológicos sería el objeto resultante de la suma de una mesa y un vaso.

to express” (2008, p. 387). En este sentido, Hirsch sostiene que los lenguajes ontológicos son *equivalentes* (recordemos (i)), por lo que podremos decir que ninguno de ellos está *expresivamente empobrecido* en términos de sus descripciones de la realidad. Me referiré a esta idea con la etiqueta “*equivalencia expresiva* de los lenguajes ontológicos”.

El problema

Ahora bien, dado el marco anterior surge un problema, a saber, que Hirsch acepta un principio conocido como **(T-weak)**, cuyo origen es una versión debilitada del **(T-strong)** de Eklund. Según Eklund: “(T) For a sentence of the form “F(a)”, of any language, to be true, the singular term “a” must refer” (2009, p. 145). Según Hirsch (2009, p. 249), la formulación adecuada de **(T-weak)** debería ser algo similar a esto: *For a sentence of the form “F(a)”, of this language, to be true, the singular term “a” must refer*. Es decir, en cada lenguaje, sus respectivos términos singulares deben tener referencia dentro de él. Por lo tanto, tomando la oración del mereólogo del ejemplo, el término singular “mesavaso” habrá de tener referencia, pues es verdadera en el lenguaje del mereólogo.

En primer lugar, dado (i), es problemático sostener que todos los lenguajes ontológicos son equivalentes. Claramente, el lenguaje del mereólogo del ejemplo es más expresivo y, bajo la noción de Eklund, el sensitista posee un lenguaje expresivamente empobrecido. Ésto es así por el hecho de que el mereólogo refiere a más objetos en oraciones verdaderas, pues el sensitista no podría referir en una oración verdadera mediante el término “mesavaso”, debido a **(T-weak)**. Es decir, el número de oraciones verdaderas que puede expresar el mereólogo es mayor que el del sensitista, lo cual implica que el número de hechos verdaderos que el mereólogo puede describir es mayor, específicamente todos aquellos hechos verdaderos acerca de sumas mereológicas rechazadas por el sensitista. Así, (i) es rechazable, pues no todos los lenguajes ontológicos son equivalentes según la tesis de la *equivalencia expresiva* de los lenguajes ontológicos.

En cuanto a (ii), puede cuestionarse que la no equivalencia a la hora de referir a ciertos objetos mediante términos singulares no es un problema ontológico si las descripciones contrapuestas son verdaderas, cada una en su lenguaje. Es decir, que el fenómeno de poder expresar más oraciones verdaderas no es un problema ontológico. Sin embargo, ésto no es así si aceptamos **(T-weak)** y el realismo, pues mediante los términos singulares estamos refiriendo a objetos del mundo, es decir, a objetos reales. En este sentido, las oraciones verdaderas que el mereólogo puede expresar son descripciones de ciertos hechos de la realidad exclusivas de su lenguaje. Así, el rechazo de (i) por los motivos del párrafo anterior nos obligan a rechazar también (ii).

En conclusión, si tomamos el igualitarismo tal y como Hirsch lo define, entonces tenemos razones para mostrar que, dado el compromiso realista, además de la defensa de **(T-weak)**, está forzado a modificar algún componente de su teoría metaontológica, pues ni (i) ni (ii) son sostenibles a la luz de lo dicho en los párrafos superiores. En concreto,

si su objetivo pasa por sostener toda la estructura formal de su teoría, manteniendo el principio (**T-weak**) y el realismo, entonces habrá de renunciar al igualitarismo, pues las condiciones (i) y (ii) se ven violadas.

Referencias

- [1] Eklund, M. (2008). The Picture of Reality as an Amorphous Lump. En T. Sider, J. Hawthorne y D. W. Zimmerman (Eds.), *Contemporary Debates in Metaphysics*. Oxford: Blackwell Publishing.
- [2] Eklund, Matti. (2009). Carnap and Ontological Pluralism. En D. Chalmers, D. Manley, y R. Wasserman (Eds.), *Metaethaphysics. New Essays on the Foundations of Ontology*. Oxford: Oxford University Press.
- [3] Hirsch, E. (2002a). Quantifier Variance and Realism. *Philosophical Issues*, 12(1), pp. 51-73. Hirsch, E. (2002b). Against Revisionary Ontology. *Philosophical Topics*, 30(1), pp. 103-127.
- [4] Hirsch, E. (2009). Ontology and alternative languages. En D. Chalmers, D. Manley, y R. Wasserman (Eds.), *Metamethaphysics. New Essays on the Foundations of Ontology*. Oxford: Oxford University Press.
- [5] Hirsch, E. (2011). *Quantifier Variance and Realism: Essays in Metaontology*. Oxford: Oxford University Press.
- [6] Lewis, D.K. (1991). *Parts of Classes*. Oxford: Blackwell Publishing.

Abstraction's Logicality and Invariance

Ludovica Conti*
Università di Pavia

Abstract

In this talk, I aim at discussing a criterion that has been recently suggested in order to prove the logicality of the abstraction operators, when they are understood as arbitrary expressions. My double aim is to inquiry whether this criterion is sufficient to achieve this goal and to use it in order to compare second-order and first-order abstraction principles. **Keywords:** Abstraction Principles, Logicality, Invariance

1 Abstraction, Logicality and Invariance

Abstractionist theories are composed by a logical core augmented with an abstraction principle of form: $@_R\alpha = @_R\beta \leftrightarrow R(\alpha, \beta)$ – which introduces and rules an operator term-forming ($@_R$) as a new symbol of the language. Then, the logicality of these theories plainly depends on the logicality of the abstraction principles. The issue of their logicality was originally settled into the seminal abstractionist program, namely Frege's Logicism. The inconsistency of this project (i.e. a theory equivalent to second-order logic augmented with Basic Law V) seemed to determine both the inconsistency and (in a classical logic) the non-logicality of Basic Law V and – *a fortiori* – of any other abstraction principle¹.

Recently, the issue of the logicality has been resumed regarding some consistent abstraction principles (like Hume's Principle), in light of the intervening studies about logicality as permutation and isomorphism invariance (cf. [6]). In order to apply this criterion to abstraction principles, we can specify at least three different subjects to be examined: the whole abstraction principle, the abstraction relation and the abstraction operator.

Regarding the abstraction principle, the more informative criterion consists of *contextual invariance*: an abstraction principle AP is *contextually invariant* if and only if , for any abstraction function $f_R: D_2 \rightarrow D_1$ and permutation π , $\pi(f_R)$ satisfies AP whenever f_R

*ludovica.conti02@universitadipavia.it

¹We will describe a relation between the abstraction principles based on the *fineness* of their equivalence relations. Cfr. [1].

does (cf. [1]). I argue that this criterion is not adequate to state the logicality of a principle. I suggest two arguments in support of this hypothesis. Firstly, it under-determines the choice between principles that are mutually inconsistent (like Hume’s Principle and Nuisance’s Principle). Secondly, this criterion turns out to be – both formally and conceptually – dependent on constraints concerning the abstraction relation: it is provably implied by the weakest form of invariance² of the abstraction relation; furthermore, a careful review of the syntactical structure of the abstraction principles shows that the abstraction relation is the real “engine” of the abstraction principle.

Regarding the abstraction relation, we can distinguish, at least, four kind of invariance: *weak invariance*, *double invariance*, *internal invariance* and *double weak invariance* (cf. [1], [3], [4]). I briefly describe their mutual relations and emphasise that, regarding abstraction relations, a very relevant meaning of logicality is provided in terms of *internal invariance*: an equivalence relation $R(X, Y)$ is *internally invariant* if and only if, for any domain D and permutation $\pi : X \cup Y \rightarrow D$, $R(X, Y)$ if and only if $R(\pi[X], \pi[Y])$.

As anticipated, *weak invariance* of the abstraction relation is sufficient to have a *contextually invariant* abstraction principle, but none of these criteria coincides or implies the canonical notion of invariance of the abstraction operator.

1.1 Abstraction operator

Regarding the canonical reading of the abstraction operator, logicality is usually spelled out in terms of *objectual invariance*³. Such criterion fails precisely in case of operators related to invariant relations (cf. [1]).

In light of the arbitrary interpretation of the abstraction operator, a new criterion – which we will call *weak invariance* – has been proposed⁴: it consists of a generalised version of Tarskian isomorphism invariance and turns out to be satisfied, at least on some domains, by all the abstraction operators that index the equivalence classes of partitions obtained by invariant equivalence relations (cf. [8]) and by the (contextually) invariant abstraction principles. If we accepted such criterion of logicality, we would be able to classify many abstraction operators as logical symbols.

My first aim consists in further clarifying this criterion – derived (in informal terms) from the standard criterion of isomorphism invariance only by substituting the canoni-

²A relation R is *weakly invariant* if and only if, for any permutation π , $R(X, Y)$ if and only if $R(\pi[X], \pi[Y])$. Cf. [1].

³An abstraction operator $@$ is *objectually invariant* if and only if for any domain D and permutation π of D , $\pi(@_R) = @_R$ – namely, $@_R(\pi(X)) = \pi(x) \leftrightarrow @_R(X) = x$.

⁴Such criterion is available in [8] and [2]: given an isomorphism i from a domain D , let i^+ such that for every set γ of objects from D , $i^+(\gamma) = \{i(x) : x \in \gamma\}$. Then, an expression ϕ is invariant just in case, for all domains D, D' and bijections i from D to D' , the denotation of ϕ on D (ϕ^D) is such that $i^+(\phi^D) = \phi^{D'}$.

cal notion of reference with the arbitrary one. Firstly, I will formally prove that *weak invariance* of the arbitrary denotation of an abstraction operator is nothing but the contextually invariance of the corresponding abstraction principle respect to all the possible non-arbitrary denotations of the same operator. We can define all the possible precisifications of the operator on a certain domain as the ordered pairs comprising the domain and the choice of a possible (non-arbitrary) denotation of the abstraction operator. Then, we can prove that the arbitrary denotation of the operator is *weakly invariant* on a domain D if and only if every precisification on D makes the principle contextually invariant. Secondly, I will suggest that *weak invariance* does not actually depend only on the arbitrary reading of the abstraction operators but also on a specific way of modelling such sort of reference, i.e. as a collection of objects. I will recap different (e.g. semantical, epistemical and metaphysical) meanings of the notion of arbitrariness and show that only the first one seems to be plainly able to determine an invariant denotation (cf. [8]). Both the epistemical arbitrariness and the metaphysical one become able to provide the same result only by accepting a non-obvious way of modelling reference as a collection of candidate objects (cf. [2], [5]).

My second aim will consist of considering and eventually answering to some objections pointing against the adoption of such criterion – also in an adequate interpretation of the arbitrariness – as a hallmark of logicality. I will compare the satisfiability of the criterion of *weak invariance* by, on the one side, the abstraction operators and, on the other side, other variable-binding operators, like ι , ε and η (cf. [8]) – which, for brevity, we will call “choice operators”. Firstly, the abstraction operators – differently from the choice operators – are not *total*, namely they turn out to be empty whether evaluated on some domains; secondly, the logicality (weak invariance) of the abstraction operators does not coincide – differently from the logicality of the choice operators – with their purely logical definability (cf. [7]); thirdly, while the logicality of the choice operators seems to formalise a property of a whole class of similar expressions, then of the intuitive notion of choice, on the contrary, the logicality of the abstraction operators seems to regard only second-order abstraction principles, by excluding any first-order abstraction operator. Such a result will be proved as a consequence of a Tarski’s theorem about first-order predicates ([6]) and the abovementioned link between *weak invariance* of the abstraction operators and isomorphism invariance of the corresponding equivalence relations ([8]).

2 First-order Abstraction

In the last part of the talk, I will compare two schemas of, respectively, second-order and first-order abstraction principles, in order to explore whether some of the limitations

mentioned above could be overcome by the adoption of a schematic setting ([4], [8]). On the one side, a schematic second-order abstraction principle – of form $\S(EF) = \S(EG) \leftrightarrow E(F, G)$, where \S is a binary abstraction operator and E any isomorphism invariant equivalence relation – defines an abstraction function from $\wp(\wp(D) \times \wp(D)) \times \wp(D) \rightarrow D$ that satisfies the criterion of weak invariance and – differently from the specific (second-order) unary operators – is total ([8]). On the other side, I will focus a schematic first-order abstraction principle – of form $\S(Ea) = \S(Eb) \leftrightarrow E(a, b)$, where \S is a binary abstraction operator and E any first-order equivalence relation – and I will show that the abstraction function from $\wp(D \times D) \times D \rightarrow D$ that it defines is – differently from the corresponding (first-order) unary operators – total, but not isomorphism invariant.

References

- [1] Antonelli, G. A., 2010, “Notions of invariance for abstraction principles”, *Philosophia Mathematica*, 18(3): 276-292.
- [2] Boccuni, F., Woods, J., 2018, “Structuralist neologicism”, *Philosophia Mathematica*, 28(3): 296-316.
- [3] Cook, R., 2016, “Abstraction and Four Kinds of Invariance (Or: What’s So Logical About Counting)”, *Philosophia Mathematica*, 25,1: 3–25.
- [4] Fine, K., 2002, *The limits of abstraction*, Clarendon Press.
- [5] Horsten, L., 2019, “Generic structures”, *Philosophia Mathematica*, 27(3): 362-380.
- [6] Tarski, A., 1956, “The concept of truth in formalized languages”, *Logic, semantics, metamathematics*, 2 (152-278): 7.
- [7] Tennant, N., 1980, “On ε and η ”, *Analysis*, 40 (1): 5.
- [8] Woods, J., 2014, “Logical indefinites”, *Logique et Analyse*, 277-307.

Cross-Fictional Identity in the Artifactual Theory of Fiction

Matthieu Fontaine*
Universidad de Sevilla

Fictional works sometimes involve reference to real or historical characters (e.g. Napoleon, Stalin), to real places (e.g. London, Paris, Seville), and even to real events (e.g. Battle of Borodino, D-Day). They can also involve reference to fictional characters native to other fictions, as it is the case in parodies (e.g. Maurice Leblanc's Herlock Sholmès), in series (e.g. Alice, Holmes), and literary traditions and cycles (e.g. Pinocchio, Faust). Following Parsons (1980, pp 51 ff.) we can call these characters from other fictions immigrants, by contrast with native fictional characters. Fictional discourse in which it is made reference to entities that lie beyond the fiction itself is precisely what we call cross-fictional discourse. Cross-fictional discourse assumes identity preservation of characters across reality and fictional works; this is what we mean by cross-fictional identity.

We call into question this assumption from the perspective of the Artifactual Theory of Fiction (ATF); mainly the theory initiated by Thomasson (1999) who considers that fictional entities are inhabitants of the actual world just as non-fictional ones are. They are abstract artifacts, created by an author and maintained in existence thanks to the existence of copies of the original work (or the memory of the members of a cultural community). Creation is accounted for by Thomasson in relation to the phenomenological notion of ontological dependency, which forms part of the identity conditions of fictional characters. Interestingly, this explains their contribution to the meaning of sentences (apparently) referring to them. Thomasson (2003) then holds the thesis that the ATF allows a uniform account for fictional discourse, by combining *de re* pretended assertions and reference to abstract entities. Cross-fictional identity is thus easily explained, since it depends on the author's intention to refer to real or immigrant characters within his own fiction. Things are even easier if we think about the use of proper names from a Kripkean setting, where they refer to the same object in every context. Finally, an author can refer to an abstract artifact he has actually created and pretend to assert things about it in fiction; for example, that his character is a detective leaving 221b Baker Street. Since it involves an abstract artifact in the actual world and in fictional discourse, cross-fictional reference is fundamental in Thomasson's account.

However, Thomasson's proposal faces serious difficulties. First, her account is incompatible with cases in which an author has the intention to merge or split different

* fontaine.matthieu@gmail.com

characters, regardless of whether these are real or fictional. We think that a modal semantics for fictionality helps in clarifying the problem, in relation to quantification across a plurality of worlds. In this context, we see the limits of a Kripkean setting, in which we directly quantify over objects that form part of the domain of different worlds. We propose a solution in the context of Hintikka's world line semantics (WLs), in which individuals are world lines that connect different objects in different worlds. Cross-world identity is not unproblematic: World lines can merge or split, or even fail to indicate an individual in one world or another, depending on how they have been drawn. In intensional contexts, we must assume that world lines have been drawn and these world lines are the individuals we quantify over. Second, quantification over fictional characters presupposes that fictionally individuated world lines have been drawn. As stressed by Hintikka and Sandu (1995), Existential Generalization is invalidated, unless additional premises are assumed. Thus, a criterion of individuation is needed to admit fictional characters in our ontology; i.e. in the domain of discourse. Thomasson's criterion of identity does not provide a sufficient condition for the individuation of fictional characters. As a solution, we propose a syncretistic criterion according to which a character must have been created; and a world line must have been drawn between the manifestations of one and the same individual under different perspectives. In accordance with the ATF, fictional discourse can thus be explained in terms of reference to and quantification over created fictional entities. Third, we still have to explain how cross-fictional identity is possible. We argue that domains of fictions are weakly impermeable. That is, although a fiction cannot prescribe reference beyond itself, cross-fictional identity can result from interpretive efforts of the readers. More formally, this is explained by appealing to a distinction between extensional and intentional modes of predication.

References

- [1] Kripke, S. (1980). *Naming and Necessity*, Basil Blackwell, Oxford.
- [2] Hintikka, J. (1969). *Models for Modalities*, Reidel, Dordrecht.
- [3] Hintikka, J. and Sandu, G. (1995). "The fallacies of the new theory of reference", *Synthese*, vol. 104(2): pp. 245-283.
- [4] Parsons, T. (1980). *Nonexistent Objects*. Yale University Press, New Haven.
- [5] Thomasson, A. L. (1999). *Fiction and Metaphysics*, Cambridge University Press, Cambridge.
- [6] Thomasson, A. L. (2003). "Fictional Characters and Literary Practices". *British Journal of Aesthetics*, vol. 43 (2): pp. 138-157.

Diálogo en metafísica

María Dolores García-Arnaldos*
Universidad CEU-San Pablo (Madrid)

En La construcción del espacio lógico (Rayo 2013), los enunciados “solo es” (Just Is Statements) desempeñan un papel central semejante al lugar capital que ocupan los enunciados analíticos en la teoría de Carnap (1958) (Rayo 2013, 35-8). Rayo los introduce a través de ejemplos como:

- (1) Que Susan sea una hermana solo es que comparta un padre con otra persona.
- (2) Que el vaso esté lleno de agua solo es que esté lleno de H₂O.

En el caso de Carnap, la cuestión de qué enunciados aceptar como analíticos dependía de qué marco lingüístico se aceptase, y esto, a su vez, dependía de qué marco fuese más útil y productivo teniendo en cuenta sus objetivos. Así, para Carnap, qué marco lingüístico se debe aceptar es una cuestión práctica, de modo que, si la elección del marco lingüístico fuese errónea, lo sería en el sentido de que no es productivo para sus metas.

También en Rayo encontramos ese elemento práctico: los enunciados Just is que se deben aceptar dependerán de los objetivos; aunque a diferencia de Carnap, Rayo separa las cuestiones epistémicas (Rayo 2013, 35-44) y metafísicas (Rayo 2013, 56-64).

Por una parte, la cuestión epistémica que Rayo ha de resolver es cómo justificar la aceptación de un sistema de representación, además de los enunciados Just is correspondientes. Para Rayo, consiste, sobre todo, en una cuestión de “análisis de costes y beneficios” (Rayo, 2013, 37), en la que el coste de aceptar un enunciado Just is es una reducción de los recursos teóricos disponibles, y el beneficio es que se evita tener que responder a ciertas preguntas explicativas.

Por otra parte, los enunciados Just is que se aceptan constituyen el espacio lógico. Una vez definido el espacio lógico, los enunciados Just is pueden ser objetivamente verdaderos o falsos dentro del propio espacio lógico, es decir, la definición de verdadero y falso se aplica internamente. Esto plantea algunas cuestiones metafísicas: ¿cómo determinamos qué enunciados Just is elegir?, ¿existe una concepción “correcta” del espacio lógico? ¿en virtud de qué sería “correcta”? (Rayo, 2013, 56-58).

Para Rayo parecería que no hay una concepción correcta de espacio lógico. Precisamente, la noción de corrección, según Rayo, presupone una concepción de espacio lógico, por lo que no parece apropiado preguntar acerca de nuestra concepción de espacio lógico si es correcta, o de nuestros enunciados Just is, si son verdaderos.

* dolores.arnaldos@rai.usc.es

Rayo considera que, por una parte, lo que implica la elección de una concepción del espacio lógico es qué conjunto de distinciones debe desplegarse en la representación del mundo. Por otra parte, es difícil separar esto de los objetivos del sujeto que establece esa representación (Rayo, 2013, 62).

De este modo, los enunciados *Just is* se presentan como enunciados cuya verdad está totalmente determinada por el sistema de representaciones que uno ha adoptado, puesto que lo constituyen, y no son objetivamente verdaderos o falsos, sino, en el mejor de los casos, verdaderos o falsos en relación con un sistema de representación que uno adopta en función de ciertos objetivos. La postura de Rayo, así presentada por Russell (2014) es semejante al convencionalismo.

El problema del enfoque convencionalista ha sido evidenciado entre otros por Quine (1936) en "Truth by Convention". La teoría de Rayo basada en los enunciados *Just is* sufre un problema análogo: no podemos coherentemente sostener a la vez que, (i) los enunciados *Just is* se aceptan trivialmente por convención "en parte" por cómo es el mundo y que (ii) cómo es "en parte" ese mundo aceptado por convención depende de los enunciados *Just is*. Es decir, los enunciados *Just is* son presupuestos en el propio modelo trivialmente y no explicables por él.

El enfoque de Rayo presenta, además, el problema del desacuerdo, es decir, cómo podemos estar de acuerdo sobre atribuciones de conocimiento que realizamos respecto de diferentes estándares epistémicos. Podremos estar de acuerdo respecto a qué se sigue de los enunciados *Just is*, pero no respecto a qué enunciados *Just is* aceptar.

Este es un aspecto fundamental que tiene profundas implicaciones sobre la naturaleza de la verdad, la normatividad y la objetividad y que depende del punto de partida. Si el punto de partida es que las atribuciones normativas consisten en un estado complejo de aceptación de enunciados *Just is*, para estar de acuerdo respecto a qué estándar epistémico seguir, tendremos que partir de que aceptamos los mismos enunciados. De no ser así, no podremos estar de acuerdo respecto al estándar. La cuestión es: sobre qué base se puede pedir a alguien que abandone su sistema de representación a favor de otro. Si aceptamos que cada sistema de representación es tan correcto como otro, basta con que una creencia particular se fundamente en un sistema de representación con determinados enunciados *Just is* para que esté justificada. Es decir, si no hay un sistema de representación más correcto que otro, entonces no se puede pensar de modo coherente que una creencia particular no está justificada, siempre y cuando esa creencia esté basada en un sistema de representación epistémico fundamental que lo permita.

La raíz del problema es que Rayo debe comenzar justificando que no hay nada que epistémicamente privilegie un sistema de representación sobre otro. Como vemos, la postura de Rayo tiene la dificultad de justificar los enunciados *Just is* que se aceptan de modo trivial pero incondicional. Si cada sistema epistémico puede ser tan correcto como otro, si no hay un hecho objetivo correspondiente, o que el hecho de que sea o no cierto depende

del sistema de representación, o del marco lingüístico, o de la concepción del espacio lógico que uno elija aceptar (Russell, 2014) se sigue, pues, que cada proposición es tan justificable como cualquier otra.

Para Rayo (2014) no está claro cómo las cuestiones de corrección objetiva pueden ser resueltas por la forma en que el mundo es. Porque, aunque los enunciados *Just is* del sujeto en particular sean verdaderos en relación con su propia concepción del espacio lógico, la cuestión es si dicha concepción constituye un marco “objetivamente correcto” con el que evaluar la verdad de un determinado enunciado *Just is*. Y no está claro que, al adoptar una concepción del espacio lógico, uno se comprometa con una de las diferentes formas de ser el mundo, en lugar de adoptar simplemente una distinción entre formas de ser del mundo y, por tanto, con el trasfondo de una concepción del espacio lógico. No obstante, el mundo en el que vivimos es un lugar mucho más complejo y para dar cuenta de todo lo que hay que explicar, según Boghossian (2020) es probable que admitamos verdades fundamentales e irreductibles sobre la conciencia, los objetos abstractos y la modalidad, como mínimo.

Como en cualquier otro campo de la investigación, aquí nos encontramos con posibles desacuerdos entre concepciones metafísicas donde cada parte apoyará su posición con consideraciones teóricas. Argumentaremos, siguiendo a Williamson (2021) que las disputas metafísicas a partir de desacuerdos son más efectivas entendidas como diálogo que como una mera negociación metalingüística. Diálogo y disputas filosóficas implican razonamiento a través de las diferentes contribuciones. Una buena interpretación de un diálogo parte de la atribución de objetivos contradictorios de las diversas partes y, sin embargo, cada una puede esperar beneficiarse de la habilidad dialéctica de la otra, desarrollarse y conducir a un mayor conocimiento.

En la primera parte de esta presentación aclaro en qué consiste la propuesta de Rayo. En la segunda, examino con más detalle los argumentos que cuestionan el enfoque convencionalista y la crítica de Boghossian a las objeciones a la admisión de verdades fundamentales e irreductibles sobre cuestiones normativas. En la última sección, argumento que planteamientos como los de Williamson resultan beneficiosos para poner a prueba la construcción del espacio lógico y son apropiados para seguir reflexionando sobre cómo los marcos conceptuales generales y cómo los diversos enfoques respecto a la normatividad pueden ayudarnos a abordar la cuestión del desacuerdo de modo productivo.

Referencias

- [1] Boghossian, P. (2020). ‘Sociologistic Accounts of Normativity’. In Ashton, N. A., Kusch, M., McKenna, R., & Sodoma, K. A. (Eds.). *Social Epistemology and Relativism*. Routledge.
- [2] Boghossian, P. (1996). ‘Analyticity Reconsidered’. *Nous* 30: 360-391.

- [3] Boghossian, P. (2000). 'Knowledge of logic'. In Boghossian, P. & Peacocke, C. (Eds.). *New essays on the a priori*, Clarendon Press. 229-254.
- [4] Carnap, R. (1958). 'Empiricism, Semantics and Ontology'. In *Meaning and Necessity: A Study in Semantics and Modal Logic* (Appendix A), 2nd ed., 205-221. Chicago, IL and London: The University of Chicago Press.
- [5] Quine, W.V.O. (1936). 'Truth by convention'. In *Philosophical essays for Alfred North Whitehead*. New York: Longman, Green, & Company Inc.
- [6] Rayo, A. (2013). *The construction of logical space*. Oxford University Press.
- [7] Rayo, A. (2014). 'Reply to critics'. *Inquiry*, 57(4), 498-534.
- [8] Russell, G. (2014). 'Hybrid identities and just being yourself'. *Inquiry*, 57(4), 455-465.
- [9] Williamson, T. (2021). 'Disagreement in Metaphysics' (forthcoming in Maria Baghramian, J. Adam Carter and Richard Rowland (eds.), *Routledge Handbook of the Philosophy of Disagreement*, London: Routledge). (Draft of 5 February 2021). Available in: <https://www.philosophy.ox.ac.uk/files/disagreement.pdf>

Una traducción de la teoría de tipos proposicional (PTT) a la lógica multivariada de primer orden (MSL)

José Javier González López*
Universidad de Salamanca

En 1963, Leon Henkin publicó el artículo “A Theory of propositional types” (Henkin, 1963) en el que desarrollaba, de forma autocontenida, una nueva lógica (lenguaje, semántica y cálculo) llamada Teoría de tipos proposicional (en adelante, PTT) basada en considerar una familia de conjuntos, denominados dominios, de la siguiente manera: en primer lugar un dominio conformado por los valores de verdad usuales, verdadero y falso, como elementos; en segundo lugar, un dominio formado por las cuatro funciones posibles desde el primer dominio hacia sí mismo; en tercer lugar tres dominios: el de funciones posibles desde el primer hasta el segundo, el dominio de funciones desde el segundo al primero, y el dominio formado por las funciones del segundo consigo mismo; etc. Así, sucesivamente, obtenemos una familia infinita de conjuntos finitos de funciones, cuyo crecimiento y cardinalidad es descriptible explícitamente. La familia de dominios está indexada por los llamados tipos proposicionales: un conjunto infinito numerable recursivamente formado por, primero por un único elemento t que indexa el primer dominio de valores de verdad, y sucesivamente por pares ordenados donde en el primer lugar se encuentra el tipo proposicional del dominio de partida de las funciones indexadas y en el segundo el tipo proposicional del de llegada.

Lo que primero llama la atención de esta nueva lógica PTT es que con solo considerar como primitivas en nuestras expresiones a las igualdades y el abstractor lambda que nos permite construir funciones, podemos construir los restantes conectores lógicos (negación, conjunción, disyunción, implicación, etc.). No solo eso, sino que, además, construimos con ellos, de forma recursiva, todas nuestras expresiones.

Lo segundo que resaltamos es la existencia de un “Nameability Theorem” que nos permite afirmar que para cada elemento de nuestros dominios existe, al menos, una expresión, llamada expresión nombre, que nos lleva a nuestro elemento por medio de una interpretación, independientemente de la asignación de variables escogida.

Por último, tras la semántica, presentamos los axiomas, el cálculo y cómo este es completo con respecto a las reglas que le son dadas.

* josejavierngl@usal.es

La lógica destino de nuestra traducción es una lógica multivariada de primer orden (Manzano, 1996), una lógica multiversal, con las signaturas conformadas de una manera determinada. Estas contendrán un conjunto infinito numerable de variables, otro de índices, Sort, para los universos de las estructuras construido como una copia del de tipos proposicionales de PTT, un conjunto infinito numerable de símbolos no lógicos destinados, y una función que dará un elemento de Sort a cada variable y , a cada símbolo no lógico, o bien un elemento de Sort (para ser un individuo) o bien una 3-tupla para indicar que se trata de un relator (un predicado de nuestra MSL). Seguidamente, presentaremos de forma breve la semántica de MSL.

El núcleo de nuestra comunicación, la traducción, comienza exponiendo las diferentes biyecciones auxiliares que existen, por definición de cardinalidad, entre la familia de dominios PTT y la de universos de un sistema de MSL, entre los conjuntos de variables de ambas lógicas, entre la familia de dominios de PTT y los predicados de MSL, etc.

La diferencia de nuestra traducción radica en cómo se define entre expresiones: por un lado definimos una función de traducción entre expresiones de tal manera que sus interpretaciones cierren el diagrama que se conforma con la primera de las biyecciones que presentamos y, por otro lado, una segunda función de traducción entre expresiones de PTT y predicados de MSL.

Así, como los individuos de PTT son funciones, primeramente son traducidas como expresiones que se interpretan como individuos y , a la vez como relaciones entre individuos (traducción de funciones) que se comportan como imágenes de las funciones que realmente son en PTT.

Referencias

- [1] Henkin, L. (1963). A theory of propositional types. *Fundamenta Mathematicae*, 323-334.
- [2] Henkin, L. (1964). Errata to the paper “A theory of propositional types”. *Fundamenta Mathematicae*, 119.
- [3] Manzano, M. (1996). *Extensions of First Order Logic*. Cambridge University Press.

Una vindicación del conjunto universal

Alejandro Gracia Di Rienzo*
Universidad de Salamanca

La concepción iterativa de los conjuntos nos proporciona una estrategia para justificar intuitivamente la mayoría de los axiomas de la teoría de conjuntos Zermelo-Fraenkel. Esta justificación tiene dos objetivos principales: por un lado, busca disipar las posibles dudas sobre el carácter *ad hoc* de los axiomas de ZF (es decir, pretende mostrar que los axiomas no son solo un artificio pergeñado para contrarrestar las paradojas de Cantor, Russell y Burali-Forti, sino que están independientemente motivados); por otro lado, la concepción iterativa pretende establecer una suerte de imagen intuitiva del “universo conjuntista” que nos ayuda a responder ciertas preguntas que podemos plantearnos sobre los conjuntos (e. g. ¿es la relación de pertenencia bien fundada?). Es esta segunda pretensión de la concepción iterativa la que me va a ocupar aquí.

En su versión más simple (Boolos 1983, Devlin 1993, Linnebo 2017), la concepción iterativa de los conjuntos puede formularse del modo siguiente. Supongamos que queremos tener a nuestra disposición conjuntos para satisfacer nuestras necesidades matemáticas habituales (como construir los números naturales, racionales, reales, complejos, etc.). Podemos concebir los conjuntos como formados por fases (que denotaremos por “ V_α ”, para cada ordinal α). Este “proceso” de formación comienza, literalmente, *ex nihilo*: $V_0 = \emptyset$, es decir, en la primera fase no tenemos ningún conjunto generado todavía. En la fase V_1 generamos todos los conjuntos de objetos que ya estaban disponibles en la fase anterior. Es decir, sacamos el conjunto potencia de la fase 0: $V_1 = \mathfrak{P}(V_0) = \{\emptyset\}$. Después, $V_2 = \mathfrak{P}(V_1) = \mathfrak{P}(\{\emptyset\}) = \{\emptyset, \{\emptyset\}\}$. Este proceso se puede iterar indefinidamente a lo largo de los ordinales finitos (los números naturales) de tal manera que, en general, $V_{n+1} = \mathfrak{P}(V_n)$. ¿Qué hacemos en una fase indizada por un ordinal límite? Si λ es un ordinal límite, estipulamos que $V_\lambda = \bigcup_{\eta < \lambda} V_\eta$. En resumen, la generación de la jerarquía iterativa de conjuntos se puede describir así:

$$V_0 = \emptyset$$

$$V_{\alpha+1} = \mathfrak{P}(V_\alpha)$$

* a.gracia@outlook.es

$$V_\lambda = \bigcup_{\eta < \lambda} V_\eta$$

Es completamente razonable preguntarse si, entre los conjuntos generados de esta manera, hay alguno máximamente inclusivo, o sea, un conjunto cuyos elementos son todos los conjuntos.

Sabemos que no es así, ya que los axiomas de ZF que justificamos mediante esta estrategia prohíben, so pena de contradicción, que exista un conjunto universal. El argumento más simple que muestra esto es que, si existiera un conjunto universal, debería ser un elemento de sí mismo, lo que es incompatible con el axioma de Regularidad en ZF. Podemos llegar a la misma conclusión en la teoría Z, el fragmento de ZF que incluye solo los axiomas propuestos originalmente por Zermelo (1967 [1908]): supongamos, en busca de una contradicción, que existe un conjunto \mathcal{U} al cual todo pertenece. Entonces podemos obtener la siguiente instancia del esquema axiomático de Separación:

$$(i) \quad \exists x \forall y (y \in x \leftrightarrow (y \in \mathcal{U} \wedge y \notin y))$$

Pero puesto que todo pertenece a \mathcal{U} , (i) es equivalente a:

$$(ii) \quad \exists x \forall y (y \in x \leftrightarrow y \notin y)$$

Esta sentencia es notoriamente contradictoria (es la paradoja de Russell). Así que negamos el supuesto: no existe el conjunto \mathcal{U} .

Es importante tener en cuenta que estas dos pruebas dependen esencialmente del axioma de Regularidad y del esquema de Separación, que son el sello característico de ZF. Así que hay que tener cuidado de no apresurarse a concluir que la noción de “conjunto universal” es intrínsecamente contradictoria. Es incompatible con los axiomas de ZF, sí, pero esto solo quiere decir que la noción de “conjunto universal” es contradictoria si pensamos que la concepción iterativa de los conjuntos es la única concepción viable, de tal manera que si un conjunto queda excluido por tal teoría es que *realmente no es un conjunto*. Esto es lo que me dispongo a refutar.

La teoría de conjuntos más célebre por preservar el conjunto universal como un conjunto *bona fide* es el sistema “New Foundations” (NF) de Quine (1937). Un rasgo atractivo de NF es que admite una formulación muy escueta: se basa únicamente en un

axioma y un esquema axiomático¹. El axioma es el de Extensionalidad, que comparte con ZF:

$$(EXT) \forall x \forall y (\forall z (z \in x \leftrightarrow z \in y) \rightarrow x = y)$$

El esquema axiomático es el esquema de Comprensión Estratificada:

$$(E-COMP) \exists x \forall y (y \in x \leftrightarrow \phi) \text{ [donde } \phi \text{ es una fórmula estratificada}^2 \text{ en la que la variable } x \text{ no está libre]}$$

Para ver que NF sí admite el conjunto \mathcal{U} basta con observar que “ $y = y$ ” es una fórmula estratificada, así que tenemos la siguiente instancia de (E-COMP):

$$(iii) \quad \exists x \forall y (y \in x \leftrightarrow y = y)$$

Puesto que, por lógica, todo es idéntico a sí mismo, esta condición resulta ser equivalente a:

$$(iv) \quad \exists x \forall y (y \in x)$$

Es decir, existe un conjunto al cual todo pertenece (\mathcal{U}). Además, las deducciones usuales de las paradojas de Cantor, Russell y Burali-Forti quedan obstruidas en NF al igual que en ZF. Esto ya debería hacernos pensar que la noción de conjunto universal es *prima facie* coherente, ya que es admisible en una teoría de conjuntos que es *prima facie* consistente.

Contra esto suele aducirse la objeción de que NF y teorías afines como ML (también propuesta por Quine (1981)) carecen de un trasfondo intuitivo como el que la concepción iterativa proporciona para ZF. Linnebo (2006:156-157) plantea claramente la objeción³:

[...] all known set theories with a universal set, such as Quine’s New Foundations, are not only technically unappealing but have lacked any satisfactory intuitive model or conception of the entities in question.

Creo que frente a esta objeción caben al menos dos réplicas.

En primer lugar, en lo que respecta a la motivación intuitiva de NF, podemos establecer el siguiente condicional: *si* la teoría simple de tipos (TST) está intuitivamente motivada, *entonces* NF también lo está. La razón de esto es que NF puede obtenerse

¹ Cabe mencionar que, a diferencia de ZF, NF es finitamente axiomatizable (Hailperin 1944).

² Para la definición de “estratificación”, véase Quine (1937: 78).

³ Boolos (1983) también la plantea en términos muy semejantes.

directamente a partir de TST (que a su vez es una simplificación del sistema de *Principia Mathematica*) mediante una modificación relativamente superficial (Forster 2018)⁴. Pero NF no solo hereda el trasfondo intuitivo de TST, sino que lo supera. Por ejemplo, NF permite construir los cardinales de Frege-Russell de una manera mucho más natural que TST: en NF hay *un* conjunto único que es el número 3, a saber, el conjunto de todos los conjuntos con tres elementos. Lo mismo no puede decirse de TST: ahí no hay tal cosa como “el” número 3, sino que existe un “número 3” en cada nivel de la jerarquía de tipos. Obviamente, que este argumento sea efectivo hay que establecer que la teoría simple de tipos está intuitivamente motivada. Manzano (1996: 181-182) recoge algunas consideraciones que apoyan esta idea.

En segundo lugar (y este es el punto más importante que quiero plantear) podemos motivar NF si la concebimos como una elucidación de una noción de “conjunto” que ya está implícita en nuestra comprensión pre-teórica de los conjuntos. Existe una tensión entre dos aspectos del concepto pre-teórico de conjunto. Por un lado, un conjunto puede entenderse como una *colección*, algo así como el resultado de una operación ideal de reunir varios objetos en una unidad. Es esta idea la que está a la base de la concepción iterativa, como sugiere el lenguaje dinámico de “formación de conjuntos” que se suele usar para formular dicha concepción. Podemos ver ZF como un desarrollo riguroso de esta concepción de los conjuntos. Por otro lado, un conjunto puede verse como la extensión de un concepto. Y es esta concepción de los conjuntos la que, a mi juicio, está a la base de NF.

Esta tensión a la que me refiero es la que Hao Wang expresa como la dicotomía entre el concepto “matemático” de conjunto y el concepto “lógico” (1983: 537), dicotomía que atribuye a Gödel. En realidad, la tensión se puede apreciar ya en la definición informal de “conjunto” que ofreció Cantor (2006: 136):

[...] entiendo en general por variedad o conjunto toda multiplicidad que puede ser pensada como una unidad, esto es, toda colección de elementos determinados que pueden ser unidos en una totalidad mediante una ley. Y creo haber definido con ello algo semejante al εἶδος o ἰδέα platónicos.

⁴ Básicamente, la noción de “fórmula estratificada” es equivalente a la noción de fórmula bien formada en TST. Lo que hace Quine al formular NF es liberalizar el lenguaje de TST, admitiendo fórmulas que son absurdas en TST (como “ $x \in x$ ”), pero estipulando que solo aquellas fórmulas admisibles en TST se puedan usar para definir conjuntos.

Al decir que un conjunto es una multiplicidad que podemos pensar como una unidad, Cantor está insinuando la concepción de los conjuntos como colecciones, mientras que al hacer alusión a una “ley” y a las ideas platónicas parece estar sugiriendo la concepción lógica de los conjuntos⁵.

Para reforzar esta idea, consideremos cómo plantea Shapiro la dicotomía entre la noción iterativa y la noción lógica de conjunto (Shapiro 1999:59):

In every context, logical sets have a Boolean structure, in that every set S has a complement, the collection of objects that are not in S . There is also a universal set, which is the complement of the empty set. In contrast, iterative sets are not Boolean, since the complement of an iterative set is not an iterative set.

Esto que dice Shapiro es muy significativo ya que, en efecto, el universo de NF tiene una estructura booleana: todo conjunto tiene un complemento y todo par de conjuntos tiene una intersección y una unión (cosa que no ocurre en ZF). Esto nos proporciona una justificación para ver NF como una elucidación de la noción lógica de conjunto, en contraste con ZF, que elucida la noción iterativa. Queda así respondida la objeción de Linnebo, y vindicada la noción de conjunto universal. (Con “vindicada” no quiero decir que *exista* un conjunto universal. Tal afirmación va en contra del enfoque pluralista que estoy sugiriendo. La noción de conjunto universal queda vindicada en la medida en que explicamos por qué es admisible en una teoría como NF y no en una como ZF: son teorías que elucidan distintas concepciones de lo que es ser un conjunto. Según la primera concepción, la noción de conjunto universal es perfectamente legítima, pero no según la segunda).

Para terminar, me gustaría observar que esta interpretación que he propuesto de la diferencia entre NF y ZF es similar a la que planteó Carnap (1945) acerca del concepto de probabilidad. Según Carnap, no hay solo un concepto de probabilidad, sino al menos dos (la probabilidad entendida como grado de confirmación y como límite de frecuencias relativas). Con esta suerte de desambiguación, Carnap supo dar cuenta de la pluralidad de teorías filosóficas de la probabilidad. Según él, esas distintas teorías no son elucidaciones rivales de un mismo concepto, sino elucidaciones de conceptos diferentes. Esto es precisamente lo que ocurre, a mi juicio, con el concepto de conjunto. ZF y NF no son dos

⁵ Esta interpretación también la ha sugerido José Ferreirós (Cantor 2006: 155).

elucidaciones de “el concepto de conjunto” entre las que haya que elegir. Son dos elucidaciones de conceptos distintos. Según uno de esos conceptos, no puede haber tal cosa como un conjunto universal. Pero según el otro, tal objeto es perfectamente admisible.

Referencias

- [1] Boolos, G. (1983). “The iterative concept of set”, en Putnam, H., Benacerraf, P., *Philosophy of Mathematics*, pp. 486-502. Cambridge, Cambridge University Press.
- [2] Cantor, G. (2006). [1883]. *Fundamentos para una teoría general de conjuntos*. Barcelona, Crítica. (Trad. José Ferreirós).
- [3] Carnap, R. (1945). “The Two Concepts of Probability”, *Philosophy and Phenomenological Research*, 5:4, pp. 513-532.
- [4] Devlin, K. (1993). *The Joy of Sets*. Nueva York, Springer.
- [5] Forster, T. (2018). “Quine’s New Foundations”, *The Stanford Encyclopedia of Philosophy* (Summer 2019 Edition), Edward Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2019/entries/quine-nf/>
- [6] Hailperin, T. (1944). “A Set of Axioms for Logic”, *The Journal of Symbolic Logic*, 9:1, pp. 1-19.
- [7] Linnebo, Ø. (2017). “Sets, Properties, and Unrestricted Quantification”, en Rayo, A., Uzquiano, G., *Absolute Generality*, pp. 149-178.
- [8] Manzano, M. (1996). *Extensions of First Order Logic*. Cambridge, Cambridge University Press.
- [9] Quine, W. V. O. (1937). “New Foundations for Mathematical Logic”, *American Mathematical Monthly*, 44, pp. 70-80.
- [10] Quine, W. V. O. (1981). *Mathematical Logic* (2^a ed.). Cambridge (MA), Harvard University Press.
- [11] Shapiro, S. (1999). “Do not claim too much: Second-order Logic and First-order Logic”, *Philosophia Mathematica*, 3:7, pp. 42-64.
- [12] Wang, H. (1983). “The concept of set”, en Putnam, H., Benacerraf, P., *Philosophy of Mathematics*, pp.530-570. Cambridge, Cambridge University Press.
- [13] Zermelo, E. (1967). [1908]. “Investigations in the foundations of set theory I”, en van Heijenoort, J. (Ed.), *From Frege to Gödel*. Cambridge, Harvard University Press, pp. 199-215.

Constitutional logics as hyperintensional representations of extensional logics

Peter Verdeé*

Universithé Catholique de Louvain

Pierre Sain-Germier**

Universithé Catholique de Louvain

Joao Daniel Dantas de Oliveira***

Universithé Catholique de Louvain

Peter Verdeé, Pierre Sain-Germier, J. Daniel Dantas de Oliveira y

Pilar Terrés Villalonga****

Universithé Catholique de Louvain

In this talk, we present a general way to represent truth-functional logics hyperintensionally. We will do this by means of the concept of a *constitution*. This is a series of meaning postulates for formulas of the language. Together they completely define the logic qua (ordinary) relation of truth preservation, but on top of that, they also constitute a theory of *logical ground* (in the sense of [3]), which allows us to define other hyperintensional relations such as exact truthmaking, hyperintensional content, and relevant entailment in a unified way.

Constitutions are sets of sequents, one formula of which is the *groundee*, the other ones are the *grounds*. We only underline the groundee. The members of a constitution (i.e. the *constitutional sequents*) that are of the form $A_1, \dots, A_n \triangleright \underline{C}, B_1, \dots, B_m$ can be interpreted as stipulating that the truth of A_1 to A_n together with the falsity of B_1 to B_m grounds the truth of C . Constitutional sequents of the form $D_1, \dots, D_n, \underline{F} \triangleright E_1, \dots, E_m$ can be interpreted as stipulating that the truth of D_1 to D_n together with the falsity of E_1 to E_m grounds the falsity of F . We can then conceive of a transitive global relation of partial ground between pairs of a formula and its truth value. For sets of such sequents to count as viable constitutions they need to establish a relation of partial ground that is well-founded on the set of all pairs of a formula and a truth value. This way we guarantee that truth values are determined in a grounded way based on the truth values of semantically less complex formulas.

* peter.verdee@uclouvain.be

** pierre.saint@uclouvain.be

*** joao.dantas@uclouvain.be

**** pilar.terres@gmail.com

Many traditional truth-functional logics can be defined by means of their constitution. As an example consider the most natural constitution of classical logic.

$$\begin{array}{c|c|c} \frac{\triangleright A, \underline{A} \rightarrow B}{A \triangleright \underline{A} \vee B} & \frac{B \triangleright \underline{A} \rightarrow B}{B \triangleright \underline{A} \vee B} & \frac{A \rightarrow B, A \triangleright B}{\underline{A} \vee B \triangleright A, B} \\ \frac{A, B \triangleright \underline{A} \wedge B}{A, B \triangleright \underline{A} \wedge B} & \frac{\underline{A} \wedge B \triangleright A}{A, \underline{\neg A} \triangleright} & \frac{\underline{A} \wedge B \triangleright B}{A, \underline{\neg A} \triangleright} \\ \frac{\triangleright A, \underline{\neg A}}{A, \underline{\neg A} \triangleright} & & \end{array}$$

If the second schema for negation ($A, \underline{\neg A} \triangleright$) is removed one gets a constitution for Batens's paraconsistent logic **CLuN** and if $\triangleright A, \underline{\neg A}$ is removed, one gets a constitution for Batens's paracomplete **CLaN**. If both are removed one gets Batens's **CLoN**. If constitutional sequents that express De Morgan and Double Negation reductions are added to respectively the constitutions of **CLuN**, **CLaN** and **CLoN** (cf. [2]), one respectively gets constitutions of **LP**, **K3** and **FDE**. Moreover, constitutions can generically be developed from deterministic and indeterministic finite matrices of many-valued logics [1]. They can also be automatically obtained from ordinary cut-eliminable analytic sequent calculi.

Such a constitution is a syntactical object that could be said to completely define a language's compositional semantics. However, they are compatible with several consequence relations that may vary according to the acceptance or rejection of additional structural rules:

- Primitive identity. $\frac{\emptyset}{\underline{A} \triangleright \underline{A}} \text{ PrId}$ where A is primitive.
- Mixed cut. $\frac{\Gamma_1, \underline{A} \triangleright \Delta_1 \quad \Gamma_2 \triangleright \Delta_2, \underline{A}}{\Gamma_1, \Gamma_2 \triangleright \Delta_1, \Delta_2} \text{ MC1}$ and $\frac{\Gamma_1 \triangleright \Delta_1, \underline{A} \quad \Gamma_2, \underline{A} \triangleright \Delta_2}{\Gamma_1, \Gamma_2 \triangleright \Delta_1, \Delta_2} \text{ MC2}$
- Full identity. $\frac{\emptyset}{\underline{A} \triangleright \underline{A}} \text{ Id}$
- Full cut. $\frac{\Gamma_1, \underline{A} \triangleright \Delta_1 \quad \Gamma_2 \triangleright \Delta_2, \underline{A}}{\Gamma_1, \Gamma_2 \triangleright \Delta_1, \Delta_2} \text{ FC}$
- Weakening. $\frac{\underline{\Gamma_1} \triangleright \underline{\Delta_1}}{\underline{\Gamma_1}, \underline{\Gamma_2} \triangleright \underline{\Delta_1}, \underline{\Delta_2}} \text{ We}$

If we close the constitutions under all of those rules and then only look at grounded sequents (i.e. sequents in which all formulas are underlined), we of course get the ordinary *Tarski consequence relation* (one that is monotonic, reflexive and transitive) of the logic (of the systems mentioned above, for example). However, already closing under just Mixed Cut and only looking at sequents with just one underlined formula gives a very interesting consequence relation: it gives a formal theory of logical ground. Adding Identity to the latter, one gets a theory for the novel concept of *loose ground*, i.e. a concept

of ground where parts of formulas that are true or false by mere logic are not required as grounds, e.g. in classical logic p loosely grounds $p \wedge (q \vee \neg q)$.

A noteworthy consequence relation is obtained by closing the constitutions under the two first rules above and only looking at the resulting grounded sequents: a rich (non-monotonic and non-transitive) consequence relation called *relevant entailment*. This relation is *relevant* in the sense that Weakening is not admissible and that valid sequents hold in virtue of links between the grounds of each of the formulas involved. For ordinary constitutions (ones in which formulas are never grounded on formulas that contain other variables than those occurring in the groundee), this entails that the consequence relation has the *variable sharing property*: each formula shares at least one variable with at least one of the other formulas in the sequent. The logic is relevant in that sense, but it is conceptually not weaker than the ordinary irrelevant Tarski logic; in many cases (all of the systems mentioned above) the existing Tarski logic is already obtained by simply closing relevant entailment under Weakening.

On top of the flexibility to both define ordinary extensional relations and new interesting hyperintensional relations, without changing the meaning of formulas (in some sense), the constitutions also allow us to develop finer-grained semantics for the languages under consideration. We will show that each constitutional language can generically be given an exact truthmaker semantics (as first proposed by Van Fraassen, cf. [5], and later elaborated by Fine, cf. [4]). The truthmaker semantics is sound and complete w.r.t. the closure of constitutions under PrId and MC. Once we have the truthmaker semantics, even more interesting logical concepts can be explicated generically : (*hyperintensional*) *content*, *exact consequence*, Parry-style *analytic implication*, and *logical synonymy*. At the same time the ordinary extensional concept of *truth in a world* can be defined, which gives us the original Tarski logic again when considering mere preservation of truth-in-a-world).

One can interpret this combination of extensional (based on truth functions and the truth-preservation relation they define) and hyperintensional approaches (based on relations between exact verifiers and falsifiers) in a pluralistic and pragmatic way: one could say that there is no conflict between the extensional and the hyperintensional viewpoints. Given a same language with the same logical meaning, the ordinary Tarski consequence relation defines the correct relation of truth preservation. Relations such as logical grounding, meaning containment, logical synonymy, and relevant entailment are hyperintensional relations one can define with respect to *the same language* with the same meaning. A logic then is a complex body of, not only a formal language and a consequence relation, but also more fine-grained aspects of the meaning of formulas in the language. It depends on what one wants to do with that language, whether one is able to restrict oneself to a coarse-grained concept like truth preservation, falsity avoidance, theoremhood or log-

ical falsehood or whether the usage requires more fine-grained concepts like grounding or synonymy. The crucial idea is here that there is no contradiction between endorsing an extensional concept of truth-in-a-world and still claiming that the full meaning of (the truth of) a formulas can only be captured by also looking at the exact states of affairs that make formulas true. Tautologies, for example, can be defended as true in all worlds, but nevertheless ‘it either snows or it doesn’t’ and ‘if John is married, then he is also married or homeless’ are made exactly true by very different states of affairs (notice that a state of not snowing and John’s state of being married or homeless concern completely different subject matters). It is then perfectly reasonable to consider that containment of truthmakers is different from preserving truth, but neither the first nor the second is mistaken.

One could also interpret our formal tools in a revisionary manner. One might argue that logic is fundamentally only about validity and that this concept cannot properly be captured correctly without making enough distinctions between the meanings of formulas. Traditional logic’s account of validity cannot distinguish between fundamentally different propositions such as tautologies that are about completely different subject matters. Traditional Tarski logics thus make the mistake of equivocating meanings that are fundamentally different. Moreover many traditional logics commit fallacies of relevance. They allow concluding arbitrary formulas from formulas that cannot be true in any model and they allow concluding formulas that are always (trivially) true from arbitrary formulas. The relevant entailment relation given here does not validate these fallacies. The framework nevertheless provides an error theory that explains where the classical consequence relations went wrong (i.e. in allowing unrestricted Weakening syntactically and in not taking into account exact truthmakers but just the much more blunt concept of truth-in-a-world).

In the talk we introduce the constitutions, give the most prominent examples, and provide the truthmaker semantics. We will show and shortly motivate the different extensional and hyperintensional logical concepts that can be defined by means of the semantics.

References

- [1] Arnon Avron and Anna Zamansky. Non-deterministic Semantics for Logical Systems. In *Handbook of Philosophical Logic*, pages 227–304. Springer, 2011.
- [2] Diderik Batens, Kristof De Clercq, and Natasha Kurtonina. Embedding and interpolation for some paralogics. the propositional case. *Reports on Mathematical Logic*, 33:29–44, 1999.

- [3] Fabrice Correia. Logical grounds. *Review of Symbolic Logic*, 7(1):1–29, 2013.
- [4] Kit Fine. Truthmaker Semantics. In Bob Hale, Crispin Wright, and Alexander Miller, editors, *A Companion to the Philosophy of Language*, chapter 22, pages 556–577. Wiley Blackwell, second edition, 2017.
- [5] Bas C. Van Fraassen. Facts and tautological entailments. *Journal of Philosophy*, 66(15):477–487, 1969.

The Final Cut

Elia Zardini*
Complutense University of Madrid

In a series of works, P. Cobreros, P. Égré, D. Ripley and R. van Rooij have proposed a nontransitive system (call it ‘K3LP’) as a basis for a solution to the semantic paradoxes. I critically consider that proposal at three levels. At the level of the background logic, I present a conception of classical logic on which K3LP fails to vindicate classical logic not only in terms of structural principles, but also in terms of operational ones. At the level of the theory of truth, I raise a cluster of related philosophical difficulties for a K3LP-based system of naive truth. At the level of the theory of validity, I consider an extension of the K3LP-based system of naive validity that is supposed to certify that validity in that system does not fall short of naive validity, and argue that such an extension is untenable.

* eliazardini@yahoo.it

FILOSOFÍA DE LA MENTE
Y EPISTEMOLOGÍA

Teorías de la conspiración: explotación, producción y tergiversación de injusticias epistémicas

Daniel Barbarrusa*
Universidad de Sevilla

Lola Medina Vizquete**
Universidad de Sevilla

Resumen

Según Coady (2019), el término “*conspiracy theories*” se utiliza para sostener el orden vigente, denostando a críticas al poder sostenidas por sectores excluidos. Sin embargo, no toda reivindicación de una identidad excluida tiene un carácter transformador ni progresista (Harding, 2006). Rechazamos la idea de Coady proponiendo la unión de los campos de teorías de la conspiración e injusticias epistémicas.

Considerando los motivos epistémicos, existenciales y sociales para creer en las teorías conspirativas (Douglas, Sutton, & Cichocka, 2017), defendemos que las personas excluidas, que sufren injusticias epistémicas (Fricker, 2007) son especialmente susceptibles de aceptar estas ideas, que prometerían soluciones y explicaciones erradas a sus problemas. Así, las teorías de la conspiración se cebarían con una población vulnerable y la dañarían aún más al privarla de explicaciones más acertadas para comprender y transformar su propia situación. Finalmente, también negamos que las “teorías de la conspiración” sufran una injusticia testimonial.

Propuesta de comunicación

El gobierno estadounidense utiliza el VIH para asesinar a los afroamericanos. Los detalles cambian, pero puede que el virus fuera diseñado en un laboratorio, o más siniestro aún, que los médicos contagien a propósito a bebés. Esta es la conclusión a la que ha llegado parte de la población afroamericana según distintos estudios (Crocker *et al*, 1999). Después de una larga experiencia de abusos y desencuentros con el gobierno, algunas personas llegaron a cuestionar la “versión oficial” acerca de la procedencia desconocida del VIH y atribuirle orígenes más ominosos. Huelga decir que su situación no mejoró después de este “descubrimiento”, sino que además desarrollaron hábitos dañinos al desconfiar de las advertencias gubernamentales contra el SIDA.

* barbarrusa@us.es

** lmvizquete@us.es

Según Annesley (2020), muchísimas mujeres han sido ignoradas en tantas ocasiones a la hora de explicar sus propios síntomas (como el dolor o el cansancio) por parte del personal sanitario, que han llegado a tener motivos para desconfiar de las soluciones propuestas por estos. En consecuencia, estas mujeres pueden buscar información en medios informales y, animadas por la sospecha hacia el sistema médico, ser más receptivas a las explicaciones sobre titánicas maquinaciones que perpetrarían el sistema médico y farmacéutico. De nuevo, esta creencia podría llevarlas a hábitos peligrosos o caer en manos de los emprendedores de las pseudo-terapias.

En ambos casos, al tratar de solucionar las injusticias que sufrían, estas personas toparon con una explicación conspiracionista que las dejaba en una situación todavía peor que el punto de partida. Ejemplos así hacen pertinente unir el estudio de las teorías de la conspiración con las injusticias epistémicas. Apostamos por defender que las teorías conspirativas pueden explotar la situación de injusticia hermenéutica de comunidades excluidas para conseguir aumentar el número de adeptos en sus filas.

Hasta donde sabemos, existen pocos textos que relacionen ambas nociones. Nuestro argumento proporciona así, por un lado, un nuevo campo de aplicación a las injusticias epistémicas, a la vez que permite dar cuenta de un tipo de injusticia que aún no ha sido explorado. Consideramos, además, que identificar esta injusticia es especialmente relevante en el contexto contemporáneo.

Para las teorías de la conspiración, por el otro lado, la idea de injusticia epistémica es relevante para estudiar cómo estas creencias pueden explotar la situación de vulnerabilidad de algunos grupos. Además, afinar el análisis previene que los propios teóricos de la conspiración aleguen estar sufriendo una injusticia cuando se desacredite su discurso, como de hecho intentan.

De esta manera defendemos una tesis positiva y otra negativa. En primer lugar, sostenemos que las teorías de la conspiración explotan injusticias hermenéuticas existentes, con la consecuencia de producir nuevas injusticias epistémicas sobre la comunidad que acepta la teoría. Estos dos momentos —explotación y producción de injusticias— podrían formar un círculo vicioso. Por otro lado, negamos que las teorías conspiranoicas, al ser categorizadas como tales, sufran una injusticia testimonial.

Según Coady (2019), el término *conspiracy theories* se utiliza para sostener el orden vigente, denostando con frecuencia a críticas al poder sostenidas por sectores excluidos. El término “teoría de la conspiración” arrastraría un *prejuicio* por el que señalamos *cualquier* denuncia de una conspiración por parte del poder como irracional. Por consiguiente, Coady concluye que deberíamos dejar de usarlo.

Asumiendo la definición que el mismo Coady señala como “cargada” de las teorías de la conspiración nos distanciamos de su propio punto de vista. Mientras que él cree que la categoría “teoría de la conspiración” se aplica (para mal) sobre cualquier denuncia, nosotras sólo nos centraremos en las teorías de la conspiración que sí son claramente irracionales. Se trata, por tanto, de emplear el término “teorías de la conspiración” como una

evaluación consciente de un grupo de ideas, no como una categoría estigmatizada debido a prejuicios. Dicho esto, para evitar confusiones entre un uso y otro, en el resto del texto nos referiremos al grupo de teorías irracionales como conspiranoicas.

Usaremos la distinción de Goetze (2018) de hasta seis tipos de injusticias hermenéuticas, a partir del desarrollo del concepto de “disentimiento hermenéutico”. En particular, consideramos que sujetos que sufren dos de los seis tipos de injusticias hermenéuticas propuestas por el autor, pueden ser más proclives a sostener teorías conspiracionistas. Estas injusticias serían las de borrado hermenéutico y de separación hermenéutica. Estos dos casos son los que interesan a nuestros propósitos ya que son los únicos en los que el propio sujeto que sufre la injusticia carece de herramientas interpretativas, y por tanto podría acudir a las ideas conspiranoicas para suplir esa carencia. En los casos restantes, el agente excluido sí tiene las herramientas para dar cuenta de su experiencia, mientras que es otro grupo el que no comprende la situación del agente. Por consiguiente, en estos casos el agente excluido no tiene un vacío conceptual que la idea conspiranoica pueda rellenar.

Además, debemos tener en cuenta que los motivos para creer en una teoría conspirativa no son meramente epistémicos (describir la realidad). Según Douglas, Sutton y Cichocka (2017), también debemos considerar motivos existenciales (como sentirse en control de la situación) y sociales (como formar parte de una comunidad).

La combinación de las dos aportaciones —la existencia de injusticias hermenéuticas y las diferentes motivaciones para aceptar ideas conspiranoicas— facilita ver cómo las teorías de la conspiración podrían explotar la carencia en estas comunidades de recursos para comprender sus propias experiencias, prometiéndoles una satisfacción a los tres grupos de necesidades: ofrecerles una explicación (con una hipótesis central simplista y que señale culpables) de su situación, que les sirva para sentirse en control de lo que sucede, y para agruparse en torno a unas ideas.

Esta promesa, en realidad, perjudica a la comunidad que la acepta. Recordemos que nos centramos en el caso de creencias conspiranoicas (irracionales), por lo que establecemos la explicación que ofrecen como equivocada por definición. La creencia conspiranoica resultaría así muy dañina al privar a la comunidad de explicaciones más acertadas para comprender y transformar su propia situación; además de que podrían estigmatizar aún más al colectivo y favorecer que él mismo se distancie del resto.

Dado que una de las motivaciones de Coady para desaconsejar el uso del término de “teorías de la conspiración” es su preocupación por que puedan fomentar injusticias de tipo testimonial (Coady, 2019, 173), creemos importante aclarar cómo, por parte de quién y sobre quién se produce el daño epistémico en este contexto. En particular, rechazamos su idea de que las teorías señaladas como conspiracionistas sufran dicha injusticia. La injusticia testimonial aparece cuando el testimonio de un grupo recibe menor credibilidad de la que merecería (Fricker, 2007)¹; sin embargo, la injusticia no tiene lugar cuando se

¹ En caso de haber injusticia, podría ser por exceso de credibilidad (Medina, 2013) (Lackey, 2018).

trata de un testimonio que no amerita credibilidad. Este es el caso cuando el testimonio constituye una mala explicación; o, como puede suceder en las creencias conspiranoicas, si ni siquiera existe interés por elaborar una explicación coherente. Es plausible que estas se limiten a crear problemas o promover una agenda política.

Como conclusión, consideramos que utilizar este lenguaje de conspiración de la manera cargada (o con términos que eviten la confusión, como conspiranoia, conspiracionismo, y derivados) resulta útil para una tesis positiva y otra negativa.

En la positiva identificamos dos procesos: que las ideas conspiranoicas explotan las injusticias hermenéuticas que sufren algunos grupos, con el fin de ganar adeptos (1) y como consecuencia, estas ideas producen más injusticia epistémica sobre el grupo al privarlos de explicaciones que permitan mayor capacidad de comprender y transformar su situación (2). Observemos, además, que al profundizar en su situación de injusticia, el proceso se podría repetir: (2) y (1) se retroalimentan.

En la tesis negativa rechazamos que las ideas conspiranoicas sufran una injusticia testimonial al ser señaladas como tal.

El mismo Coady (2019, 170) señala un grupo de teorías conspirativas como “obviamente falsas”. Utilizar un término para señalarlas es lo que nos permitirá afrontar las injusticias; de lo contrario, solamente miraríamos a otro lado mientras suceden.

Referencias

- [1] Annesley, K. (2020). Connecting epistemic injustice and justified belief in health-related conspiracies. *Ethics, Medicine and Public Health*, 15, 100545.
- [2] Coady, D. (2019). Psychology and conspiracy theories. *The Routledge handbook of applied epistemology*, 166-175.
- [3] Crocker, Jennifer; Luhtanen, Riia; Broadnax, Stephanie & Blaine, Bruce. (1999). Belief in US Government Conspiracies Against Blacks Among Black and White College Students: Powerlessness or System Blame?. *Personality and Social Psychology Bulletin*. 25. 941-953
- [4] Douglas, K. M., Sutton, R. M., & Cichocka, A. (2017). The psychology of conspiracy theories. *Current directions in psychological science*, 26(6), 538-542.
- [5] Fricker, M. (2007). *Epistemic injustice: Power and the ethics of knowing*. Oxford University Press.
- [6] Goetze, T. S. (2018). Hermeneutical dissent and the species of hermeneutical injustice. *Hypatia*, 33(1), 73-90.

- [7] Harding, S. (2006). Transformation vs. resistance identity projects: Epistemological resources for social justice movements. In *Identity politics reconsidered* (pp. 246-263). Palgrave Macmillan, New York.
- [8] Lackey, J. (2018). Credibility and the distribution of epistemic goods. In *Believing in Accordance with the Evidence* (pp. 145-168). Springer, Cham.
- [9] Medina, J. (2013). *The epistemology of resistance: Gender and racial oppression, epistemic injustice, and the social imagination*. Oxford University Press.

Alcances y límites de la noción de punto de vista desde la teoría de los espacios conceptuales

Enrico Brugnamì*
Universidad de La Laguna

En la presente comunicación me propongo analizar el concepto de Punto de Vista (PdV) desde la teoría de los espacios conceptuales de Peter Gärdenfors y Antti Hautamäki y argumentar cuales son las virtudes y los defectos de esta concepción de perspectiva. Para ello seguiré el siguiente itinerario: 1) contextualizar brevemente la teoría de los espacios conceptuales en la obra de Gärdenfors (2000, 2014); 2) ilustrar la noción de espacio conceptual desarrollada por Gärdenfors (2000) y Hautamäki (2020); 3) explicar el concepto de PdV que se extrae del concepto de espacio conceptual y 4) argumentar qué nos aporta esta manera de entender los PdV y en qué considero que falla.

En las ciencias cognitivas usualmente se ha analizado el problema del modelado de las representaciones cognitivas desde dos perspectivas: la simbólica, representada por el computacionalismo, y la asociacionista, representada por el conexionismo. La teoría de los espacios conceptuales que presenta Gärdenfors (2000) nos ofrece una nueva vía para poder hacer frente al problema de modelar representaciones sin excluir las anteriores perspectivas por completo. Así, se trata de representar información semántica a partir de estructuras geométricas en un espacio topológico concreto.

De acuerdo con Gärdenfors (2000; 2004) y especialmente con Hautamäki (2016, 2020), entendemos un espacio conceptual CS como una colección de n-dimensiones cualitativas (siendo n es un número natural igual o mayor a 0), donde son representadas tanto propiedades como conceptos. Una dimensión cualitativa es una estructura $\langle Q, D \rangle$ formada por una cualidad Q y un dominio D. Por cualidad entendemos una propiedad cualitativa de un objeto, sea esta ostensible o no (p. ej.: el color). Por dominio entendemos el conjunto de todas las posibles instanciaciones de dicha cualidad, a las cuales llamamos valores (p. ej.: rojo sería un valor en el dominio de la cualidad “color”). Dadas n-dimensiones cualitativas $\langle Q_1, D_1 \rangle \dots \langle Q_n, D_n \rangle$, generamos un espacio conceptual CS mediante el producto cartesiano de todos los conjuntos que son dominios. Es decir, $CS = \langle D_1 \times \dots \times D_n \rangle$.

Supongamos el concepto de perro. Podemos representar dicho concepto atendiendo a diversas dimensiones cualitativas, como el color del pelo, la masa corporal, la variación fenotípica, etc. Por ejemplo, un pomerania marrón con una masa corporal de 3kg y un ratero negro con una masa corporal de 10kg son dos entidades distintas que pueden

* ebrugnam@ull.edu.es

representarse en un mismo espacio conceptual simplemente variando los valores de los distintos dominios.

Las representaciones de los espacios conceptuales son sencillas cuando no superan las tres dimensiones cualitativas. El ejemplo anterior puede representarse sin dificultad en un espacio topológico vectorial, pues solo estamos considerando tres dimensiones cualitativas. Si considerásemos más, la representación vectorial no sería posible. Sin embargo, siempre podemos realizar una representación simbólica. En esta comunicación mostraremos tanto una representación vectorial de un concepto de tres dimensiones como una representación simbólica de un concepto de más de tres dimensiones.

Esta teoría de los espacios conceptuales tiene importantes consecuencias filosóficas. Una de ellas es relativa al perspectivismo filosófico. Según Liz (2013), el perspectivismo se sitúa entre el relativismo radical y el absolutismo epistémicos. La tesis principal de esta posición es que todas las cosas tienen unos aspectos dependientes de nuestro conocimiento y otros independientes. La noción de PdV o perspectiva es la que articula y equilibra esta balanza entre los modos de ser epistémicamente dependientes e independientes.

Según Hautamäki (2020), de acuerdo con la teoría de los espacios conceptuales, un PdV es un conjunto de dimensiones cualitativas seleccionadas para representar o describir un objeto según el peso de cada una. A una dimensión cualitativa se le asigna un peso relativo según la relevancia que tenga para representar conceptualmente una entidad cualquiera. La asignación del peso a las distintas dimensiones cualitativas es arbitraria y responde a fines pragmáticos. Según qué nos sea más útil en un momento determinado seleccionaremos unas dimensiones u otras. Por ejemplo, a la hora de representar geométrica o simbólicamente el concepto de perro nos serán más comunes y útiles dimensiones cualitativas como la variación fenotípica, el color del pelo o el tamaño y menos útiles otras dimensiones, como el sabor o su aerodinámica. Igualmente, aunque una representación según la variación fenotípica sea conceptual y fenoméricamente distinta de una representación según el sabor, ambas son topológicamente equivalentes. Esto es así porque no añadimos ni sustraemos dimensiones cualitativas al concepto de perro, sino que tan solo enfatizamos unas en un PdV y otras en otro PdV.

Esta concepción de los PdV tiene una principal ventaja, y es que nos permite entender y representar de forma gráfica e intuitiva las distintas variaciones que un concepto puede sufrir al cambiar la perspectiva desde la cual se lo toma en consideración, o incluso las variaciones que pueden darse dentro de una misma perspectiva. Esto es bastante útil para poder analizar casos en los que dos perspectivas parecen mostrar dos entidades completamente distintas cuando, sin embargo, lo que realmente sucede es que solo se están teniendo en cuenta distintas dimensiones cualitativas de la misma entidad.

Sin embargo, esta concepción tiene unos límites. Uno de ellos es que no siempre es posible representar gráficamente todas las dimensiones cualitativas con todos sus respectivos dominios y valores, como vemos en Hautamäki (2020). Otro límite, y este con consecuencias filosóficas importantes, es que en las representaciones de los conceptos y los

PdV se deja de lado al sujeto. Más allá de los contenidos concretos de los conceptos, las actitudes proposicionales que adoptan los sujetos epistémicos no quedan representadas. Esto deja fuera de análisis casos como el formulado en la proposición 6.43 del *Tractatus logico-philosophicus* de Wittgenstein, donde los hechos permanecen inalterados, siendo solamente el cambio en las disposiciones del sujeto lo que crea perspectivas distintas.

En conclusión, creemos que, pese a algunas de sus limitaciones, el modelo de los PdV a partir de la teoría de los espacios conceptuales de Gärdenfors y Hautamäki ofrece una herramienta muy útil y novedosa para analizar una gran variedad de casos de cambio epistémico, como cuando una perspectiva es alterada en su contenido o cuando se pasa de una perspectiva a otra para poder representar una misma entidad de manera plural.

Bibliografía

- [1] Gärdenfors, P. (2000). *Conceptual spaces. The geometry of thought*. MIT Press, Massachusetts.
- [2] — (2004). “Conceptual spaces as a framework for knowledge representation”, en *Mind and Matter*. Vol. 2 (2), 9-27.
- [3] — (2014). *The geometry of meaning: semantics based on conceptual spaces*. MIT Press, Massachusetts.
- [4] Hautamäki, A. (2016). “Points of view: a conceptual space approach” en *Foundations of Science*. Vol. 21, 493-510.
- [5] — (2020). *Viewpoint relativism. A new approach to epistemological relativism based on the concept of points of view*. Springer, Heidelberg.
- [6] Kaipainen, M. et al. (eds.) (2019). *Conceptual spaces: elaborations and applications*. Springer, Heidelberg.
- [7] Liz, M. (2013). “Analizando la noción de puntos de vista”, en M. Liz (ed.) (2013): *Puntos de vista. Una investigación filosófica*. Laertes, Barcelona. 28-164.
- [8] Wittgenstein, L. (2013). *Tractatus logico-philosophicus*. Tecnos, Madrid
- [9] Zenker, F. y Gärdenfors, P. (2015). *Applications of conceptual spaces. The case for geometric knowledge representation*. Springer, Heidelberg.

El neo-cartesianismo como callejón sin salida: cómo la psicología ecológica se acomoda mejor en las ciencias cognitivas

Joan Camarena Bononad*
Universidad de Valencia

El programa oficial de investigación de las ciencias cognitivas promueve el análisis de la cognición como procesamiento de información recibida pasivamente por los órganos sensoriales y cuyo procesamiento culmina en la formación de representaciones mentales (Thagard, 2014). Así pues, bajo esta perspectiva, el proyecto de naturalizar la intencionalidad pasa por la idea de que un estado intencional es un estado representacional con un determinado contenido el cual está sometido a unas condiciones de normatividad. Sugiero llamar a esto la Tríada RNC (Tríada Representación-Contenido-Normatividad), la cual especifica la dependencia lógica y ontológica entre estos tres conceptos. Desde una perspectiva naturalista, se necesitaría una teoría naturalista del contenido representacional (y por tanto del contenido y la normatividad) si se quiere naturalizar la intencionalidad.

Ahora bien, quien base su proyecto naturalizador sobre las representaciones internas debe, a la vez, responder a dos preguntas: la primera, ¿en virtud de qué una representación tiene *ese* contenido y no más bien otro?, y, la segunda cuestión, ¿cuáles son las condiciones por las que se distingue un estado representacional de un estado no-representacional?

Dos modelos se han posicionado como los principales a la hora de tratar estas cuestiones: por un lado las teorías informacionales, que tratan de explicar qué es una representación y cómo se determina su contenido apelando a relaciones probabilísticas (Eliasmith, 2000, 2005, 2013; Usher, 2001; Birch, 2014), y por otro lado la teleosemántica, que analiza los sistemas representacionales apelando a la noción de función (Dretske, 1981; Millikan, 1984). Estos dos modelos no son para nada excluyentes; de hecho, han llegado a combinarse entre sí para formar una teleosemántica informacional (Neander, 2017). Es más, comparten dos ideas fundamentales que subyacen en sus formulaciones: primero, que existe un contenido que se puede extraer del mundo y que los sistemas representacionales son los encargados de procesar, y segundo, que este contenido es informacional según la propuesta del procesamiento de la información de Shannon.

Sin embargo, y teniendo en cuenta lo dicho hasta ahora, debe hacerse un inciso al respecto: la idea de que los estados representacionales son estados intermedios (tal y como proponen los dos modelos explicativos anteriores) entre un sistema vivo (u organismo

* jcb121092@gmail.com

cualquiera) y el mundo (o entorno en el que habita) es una idea que encuentra sus raíces en Descartes. En este sentido, el proyecto neo-cartesiano puede definirse según la propuesta de Searle de que las propiedades intencionales dependen de propiedades semánticas y que un estado intencional representa un objeto del mundo del mismo modo que un acto de habla representa un estado de cosas del mundo (Searle, 1983). A raíz de esto, suele hacerse la siguiente inferencia: decir que un estado mental es intencional es lo mismo que decir que es representacional y que posee un contenido determinado (contenido que, según lo dicho por Searle, va a depender de propiedades semánticas tales como la referencialidad y la normatividad o condiciones de verdad).

Esto es motivo de controversia. Debido a ello, algunos autores sostienen que debe abandonarse la asunción de que la intencionalidad está *siempre* cargada de contenido (Hutto & Myin, 2013).

Así pues, tras presentar de una forma general, pero sustancial, los modelos de las teorías de la información y la teleosemántica, pasaré a mostrar cómo estos proyectos neo-cartesianos son insatisfactorios por dos motivos. El primero de ellos es metafísico, y es que las críticas a la idea de que la intencionalidad está cargada de contenido no están hechas en vano. Estas críticas, a mi modo de ver, tienen su origen en aquello que Sellars llamó “El Mito de Lo Dado”, y han evolucionado hasta nuestros días a través de la formulación del llamado “Problema Fuerte del Contenido” (Hutto & Myin, 2013). La segunda crítica al neo-cartesianismo está hecha desde el programa evolutivo en biología y desde la selección natural. Si se acepta, tal y como lo expone la Síntesis Extendida de la Evolución (Laland, Matthews & Feldman, 2016), que los organismos no son solo productos de un proceso adaptativo sometidos a presiones selectivas externas sino que juegan un papel activo en la conformación de su entorno (como se da en los procesos de construcción de nichos), se está asumiendo que las acciones y la conducta de estos influyen modificando el entorno y las presiones selectivas. Así pues, ¿qué papel juegan los estados representacionales como intermediarios entre entradas sensoriales y salidas conductuales de unos organismos que, a través de esta conducta, modifican las presiones selectivas del entorno? es decir, ¿funciona la evolución según la idea de que los estados representacionales determinan si una conducta es satisfactoria, bajo condiciones semánticas de normatividad? Mi respuesta es que no, y puede resumirse perfectamente bajo la famosa sentencia de que la selección natural no funciona según condiciones de verdad, sino sobre la base de si algo es exitosamente adaptativo o no (Stich, 1990).

Así, dado que la idea de estados representacionales que actúen intermediarios es, por un lado, metafísicamente inestable debido a su dependencia en “lo dado” y, por otro lado, biológicamente insostenible para dar cuenta de la evolución y selección natural, el siguiente y último punto en la exposición es el de encontrar un nuevo programa científicamente respetable que intente encajar bajo un nuevo marco conceptual muchas de las piezas que quedaban sueltas en el puzzle de la intencionalidad. Y este programa de estudio es la psicología ecológica.

De este modo, el trabajo concluye mostrando dos aspectos clave de la psicología ecológica que evitan los problemas del neo-cartesianismo. Por un lado, la psicología ecológica no se ve entorpecida por el Problema Fuerte del Contenido, por rechazar, entre otras cosas, el contenido informacional de Shannon (Gibson, 1979; Heras-Escribano, 2019). Así pues, uno de los modos de conclusión del trabajo es la puesta en escena de la noción de información ecológica y cómo esta puede tomar el relevo en las ciencias cognitivas contemporáneas. Por otro lado, y respondiendo a las exigencias evolutivas, se mostrará cómo gracias a las affordances (noción central en la psicología ecológica) y su relación con la construcción de nichos, el organismo puede acomodarse mejor dentro del modelo de la Síntesis Extendida adoptando un papel activo en todo el proceso evolutivo.

References

- [1] Birch, J. (2014). Propositional content in signalling systems. *Philosophical Studies* 171-3, 493-512.
- [2] Dretske, F. (1981). *Knowledge and the Flow of Information*. Cambridge: The MIT Press.
- [3] Eliasmith, C. (2000). *How neurons mean: A neurocomputational theory of representational content*. Unpublished Dissertation, Washington University in St. Louis.
- [4] Eliasmith, C. (2005). A new perspective on representational problems. *Journal of Cognitive Science* 6, 97-123.
- [5] Eliasmith, C. (2013). *How to build a brain: A neural architecture for biological cognition*. Oxford University Press.
- [6] Gibson, J. J. (1979). *The ecological approach to visual perception*. Dallas [etc.]: Houghton Mifflin.
- [7] Heras-Escribano, M. (2019). *The Philosophy of Affordances*.
- [8] Hutto, D. D., & Myin, E. (2013). *Radicalizing enactivism: Basic minds without content*. MIT Press.
- [9] Laland, K., Matthews, B., & Feldman, M. W. (2016). An introduction to niche construction theory. *Evolutionary Ecology*, 30(2), 191-202.
<https://doi.org/10.1007/s10682-016-9821-z>
- [10] Millikan, R. G. (1984). *Language, thought, and other biological categories: new foundations for realism. A Bradford book* (Nachdr.). Cambridge, Mass.: MIT Press. Retrieved from <http://files/1405/Millikan - 2009 - Language, thought, and other biological categories.pdf>

- [11] Neander, K. (2017). *A Mark of The Mental. In Defense of Informational Teleosemantics*. Cambridge: MIT Press.
- [12] Searle, J. (1983). *Intentionality*. Cambridge University Press.
- [13] Skyrms, B. (2010). *Signals: Evolution, Learning and Information*. Oxford University Press. Stich, S. (1990) *The Fragmentation of Reasons: Preface to a pragmatic Theory of Cognitive Evaluation*. Cambridge: MIT Press.
- [14] Thagard, P. (2014). Cognitive Science. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2014). Retrieved from <http://plato.stanford.edu/archives/fall2014/entries/cognitive-science/>

Contraejemplos a la solución fiabilista del problema de la inducción

Manuel Arnulfo Cañas Muñoz*
Universidad de Salamanca

1 Introducción

En décadas recientes se ha propuesto una solución al problema de la inducción desde la perspectiva fiabilista sobre la justificación epistémica y el conocimiento (Goldman, 1979, 1986). Esta respuesta aparentemente evita el escepticismo, sin rechazar completamente el argumento original de Hume. Simplemente debilita la conclusión humeana enmarcándola de una forma distinta.

Como argumenta Levin (1993), la solución fiabilista al problema de la inducción deja intacto el estatus epistémico de las inducciones cotidianas, al reconocer como un hecho que las personas ordinarias están justificadas cuando aceptan ciertas de sus creencias mediante bases inductivas. En principio, desde esta reformulación el tópico entorno a la justificación de la inducción pierde su carácter paradójico; se reduce a la pregunta respecto a si la inducción produce o no justificación o conocimiento. Sobre lo anterior el fiabilista un veredicto evidente: la inducción ofrece justificación epistémica porque es un proceso fiable. Así, si alguien no puede justificar o creer que la inducción es fiable, ello no implica que sus creencias inductivas no estén justificadas.

Además, lo anterior es coherente con otra tesis humeana sobre la inducción, el uso de la inducción no es afectado por ningún argumento en contra de la fiabilidad inductiva (Levin, 1993). Ningún elemento conceptual o matemático puede evitar que utilicemos inferencias inductivas, porque su empleo no se sostiene sobre algún tipo de razonamiento. Si la perspectiva fiabilista es la única que puede reivindicar nuestras inducciones cotidianas, entonces en apariencia esta solución es la más satisfactoria.

2 La solución fiabilista al problema de la inducción

Van Cleve (1984) y Papineau (1992, 1993) ofrecen dos claros ejemplos de la solución fiabilista mencionada; para este tipo de fiabilismo un sujeto S está justificado en creer cierta proposición si su creencia es resultado de un proceso creador de creencias fiable. Parte de la dificultad que señalaré en este ensayo radica en que la definición de fiable es

* id00771963@usal.es

demasiado ambigua, escuetamente se entiende como una mayor proporción de resultados verdaderos que falsos.

Para esta concepción, que Hume descubriera que la inducción carece de certeza no es un problema genuino, porque la fiabilidad no implica certeza. Así que, el hecho de que algunas conclusiones inductivas sean falsas no implica que la inducción no sea un método creador de creencias fiable. Pero incluso, si Hume probó que no sabemos que la inducción es fiable, eso no supone que la inducción no sea una fuente de justificación o conocimiento. Según el fiabilismo, la conclusión escéptica de Hume únicamente emerge desde una perspectiva que considere a la certeza como un requisito de la justificación epistémica. La única condición necesaria para considerar a la inducción como una inferencia que ofrece justificación es que la mayoría de sus resultados sean verdaderos. Lo anterior se satisface con una meta-inducción que apela a las conclusiones inductivas verdaderas del pasado:

Most of the inductive inferences I have drawn in the past from true premises have had true conclusions.

Hence,

The majority of all inductive inferences with true premises have true conclusions.
(Van Cleve, 1984: 557).

La objeción principal en contra de este tipo de meta-inducción que reconocen los fiabilistas mencionados es la aparente circularidad viciosa subyacente. Van Cleve y Papineau enfrentan este problema mediante la distinción entre circularidad de premisa y circularidad de regla, donde la primera es la única que se puede considerar viciosa. La circularidad de premisa sucede cuando la conclusión de un argumento está incluida dentro de sus premisas, mientras que la circularidad epistémica ocurre cuando es necesario utilizar una regla inferencial para concluir que es fiable. Dado lo anterior, la meta-inducción fiabilista aludida presumiblemente comete circularidad de regla, pero no circularidad de premisa, ya que su conclusión no forma parte de sus premisas.

Esta estrategia sigue parcialmente una solución anterior al fiabilismo que fue formulada por Braithwaite (1953), desde lo que llama un tipo de prediccionismo; una concepción que defiende que una regla se justifica a partir de su poder predictivo. Para Braithwaite, la justificación de cualquier inferencia inductiva consiste en su conformidad con las reglas inductivas, reglas predictivamente fiables, de las cuales la más importante es la inducción enumerativa; donde la aceptación de la fiabilidad de este principio principal es en sí misma una hipótesis inductiva. Braithwaite confrontó la circularidad envuelta en lo anterior de un modo similar a Van Cleve y Papineau, a través de un tipo de circularidad no viciosa llamada circularidad efectiva. En este tipo de circularidad la verdad de cierta conclusión formada mediante alguna inferencia es considerada como una condición suficiente de la corrección de dicha especie de inferencia. Braithwaite exige mucho menos

para considerar a la inducción como una fuente de justificación de lo que requiere el fiabilismo, únicamente demanda que la inducción ofrezca varias conclusiones verdaderas, y no una mayor proporción de resultados verdaderos. Sin embargo, lo anterior es posible porque para Braithwaite cualquier regla que posea fiabilidad predictiva es fundamentalmente inductiva. Esta diferencia fundamental permite que la estrategia fiabilista pueda ser empleada para justificar que la contra-inducción o la clarividencia son procesos fiables; algo que es imposible desde el prediccionismo de Braithwaite.

3 Dos posibles tipos de contraejemplos

Hay por lo menos dos modos generales de replicar la estrategia fiabilista mencionada para defender que otros procesos apócrifos no-inductivos son fiables. En primer lugar, se puede recurrir únicamente a ciertas conclusiones verdaderas ignorando las falsas, para concluir que el método en cuestión es fiable. Esto puede aplicarse a la clarividencia y a la contra-inducción:

Ejemplo α

1. En el tiempo T1, el sujeto S1 predice que X1 sucederá en el momento T2 a través de una supuesta clarividencia, y en T2, es el caso que X1.
2. En el tiempo T3, el sujeto S1 predice que X2 sucederá en el momento T4 a través de una supuesta clarividencia, y en T4, es el caso que X2.
3. Por lo tanto, la presunta clarividencia de S1 es un método fiable (por inducción).

Ejemplo β

1. En el tiempo T5, el sujeto S1 predice que X3 sucederá en el momento T6 por contra-inducción, y en T6, es el caso que X3.
2. En el tiempo T7, el sujeto S1 predice que X4 sucederá en el momento T8 por contra-inducción, y en T8, es el caso que X4.
3. Por lo tanto, la contra-inducción es un método fiable (por inducción).

La segunda manera consiste en emplear la misma inferencia para justificar su fiabilidad en lugar de recurrir a una inducción como en los dos ejemplos anteriores.

Ejemplo γ

1. La presumible clarividencia de S1 predice que la mayoría de sus predicciones serán verdaderas.
2. Por lo tanto, la clarividencia de S1 es fiable.

Ejemplo δ

1. En el tiempo T9, el sujeto S1 predice que X5 sucederá en el momento T10 por contra-inducción, y en T10, no es el caso que X5.
2. En el tiempo T11, el sujeto S1 predice que X6 sucederá en el momento T12 por contra-inducción, y en T12, no es el caso que X6.
3. Por lo tanto, la mayoría de las contra-inducciones de S1 serán verdaderas y la contra-inducción es fiable.

Una diferencia notable con el último ejemplo es que mientras más conclusiones contra-inductivas falsas se encuentran, mayor peso contra-inductivo tiene la conclusión de que los siguientes resultados serán verdaderos. Este punto se debe a la misma naturaleza apócrifa de la contra-inducción, donde las regularidades son utilizadas para inferir lo contrario. Así, la contra-inducción resulta un caso notable, no importa si hay muchas o pocas conclusiones contra-inductivas verdaderas o falsas, en cualquier situación es posible emplear de algún modo la estrategia fiabilista a favor de la fiabilidad inductiva.

4 Conclusión

Si la estrategia fiabilista discutida puede llevarnos a la conclusión de que la inducción, la clarividencia y la contra-inducción son métodos fiables, entonces algo resulta cuestionable respecto de esa solución. Aunque no es claro respecto de lo que podríamos llamar clarividencia, en lo concerniente a la contra-inducción claramente es un proceso excluyente de la inducción; ya para cualquier inducción es posible formular una contra-inducción que comparta sus premisas, pero que posea una conclusión contraria. Así, es inconsistente asumir la fiabilidad de la inducción y de la contra-inducción al mismo tiempo.

Hay algo faltante en la estrategia fiabilista discutida que evite la inconsistencia anterior, pero dado que no ofrece ningún elemento para preferir los resultados verdaderos de algún proceso creador de creencias en lugar de los falsos, ni algún método preciso para determinar la fiabilidad, esta solución es por lo menos incompleta.

5 Referencias

- [1] Braithwaite, R. (1953). *Scientific Explanation: A Study of the Function of Theory, Probability and Law in Science*. UK: Cambridge University Press.
- [2] Goldman, A. (1979). What is Justified Belief? In G. Pappas (Ed.), *Justification and Knowledge* (pp. 1-25). USA: Reidel.
- [3] Goldman, A. (1986). *Epistemology and Cognition*. USA: Harvard University Press.
- [4] Levin, M. (1993). Reliabilism and Induction. *Synthese*, 97 (3), 297-334.

- [5] Papineau, David (1992). Reliabilism, Induction and Scepticism. *Philosophical Quarterly*, 42 (166), 1-20.
- [6] Papineau, David (1993). *Philosophical Naturalism*. UK: Blackwell.
- [7] Van Cleve, J. (1984). Reliability, Justification, and the Problem of Induction. *Midwest Studies in Philosophy*, 9 (1), 555–567.

Rethinking the epistemology of deep disagreements through the lens of Wittgenstein's later philosophy

Jordi Fairhurst*
Universidad de las Islas Baleares

Consider the disagreements over whether the Earth was created by God or as a consequence of the Big Bang, or about the morality and legality of euthanasia or abortion. These disagreements are deep disagreements: they are systematic and persistent disagreements rooted in contrary worldviews where there is no mutually recognized method of resolution because the parties involved generally reason and analyze evidence using different epistemic frameworks. The systematicity of deep disagreements can be defined through the ripple-effect they cause. Deep disagreements are not isolated, but rather extend outwards, having a ripple-effect on what one is committed to disagreeing over in other cases (Ranalli 2018; Lavoriero 2018).

Wittgenstein has been at the heart of the epistemology of deep disagreements since the inception of this topic, as Fogelin (1985, 5) himself recognized. In particular, *On Certainty* (henceforth, OC) has dominated Wittgenstein's influence in the study of deep disagreements. A central theme in OC is that rational evaluation generally presupposes fixed *hinge commitments* that are the fundamental presuppositions of one's worldview, which are exempt from rational evaluation and are presupposed by our epistemic practices of giving reasons for belief and doubt.

Since the metaphor of 'hinges' plays a central role in the scholarship of OC, Wittgensteinian theories of deep disagreements (henceforth, WTDD), led by Duncan Pritchard (2009; 2011a; 2018a; 2018b), have generally characterized deep disagreements as disagreements over the different hinge commitments of one's worldview. Accordingly, WTDD have usually been based on hinge epistemology (henceforth, HE) (see Ranalli 2020; Lavorerio Forthcoming).

HE is an umbrella-term for a diverse group of epistemological theories that expand on Wittgenstein's concept 'hinges'. These theories are classified into two opposed camps as a consequence of how these hinges are conceived. Non-epistemic theories (e.g., Moyal-Sharrock 2004; 2016; Stroll 1994; Pritchard 2011a; 2011b) argue that hinges are outside the scope of rational evaluation and lack epistemic properties. Meanwhile, epistemic

* jordi-f@hotmail.com

¹ Some philosophers favor other notions, e.g., 'framework propositions' (Wright 2004b), over 'hinge commitments'. These terminological differences are not crucial for our purposes here.

theories (e.g., Kusch 2016; Williams 1991; 2004; Wright 2004a; 2004b; 2014) argue that hinges are within the scope of rational evaluation and do have epistemic properties, albeit in a non-paradigmatic way.

Defenders of WTDD, then, agree that hinges play a fundamental role in deep disagreements but disagree on whether these hinges are epistemic or non-epistemic. Thus, the commitments one undertakes in HE will have ramifications for the respective WTDD one develops. The standard approach of WTDD and HE is generally marked by putting forward an (allegedly) exclusively correct theory which claims to explain everything essential about deep disagreements and hinge commitments. Namely, it is assumed that the nature of deep disagreements and hinge commitments is epistemologically uniform and that, in consequence, they can only be adequately explained by one true global theory. This has resulted in a stalemate between competing theories, both in WTDD and HE, with no clear path to a satisfactory resolution.

I propose to develop a novel Wittgensteinian approach to deep disagreements that overcomes the limitations of this standard approach of WTDD and HE. In particular, I propose to develop a Piecemeal Wittgensteinian Epistemology of Deep Disagreements (henceforth, PWEDD) that seeks to clarify different aspects of deep disagreements and their hinge commitments by means of a variety of more local descriptions, rather than putting forward an allegedly exclusively correct theory that is said to explain all instances of these phenomena.

The central claim underling this PWEDD is that, against the standard approach of WTDD and HE, it is mistaken to assume that deep disagreements and hinge commitments are epistemologically uniform. Deep disagreements and hinge commitments are complex phenomena that cannot be reduced to a common genus of epistemological features, as it overlooks the particularities that they may have in different fields (e.g., geology, physics, ethics, etcetera). For instance, it may be the case that hinge commitments are epistemic within the context of a deep disagreement in a given field and non-epistemic in another. Every case involves certain particularities that require specific descriptions and clarifications, not global theories.

The mistake of the existing WTDD and HE theories, then, is not that they say nothing helpful about deep disagreements or hinge commitments. It is rather that each theory claims to explain everything essential about these phenomena. It is for this reason that there is a current stalemate. If, however, we rectify these theories by restricting them to local descriptions that clarify the epistemology of particular kinds of deep disagreements (and their respective hinge commitments) it seems that we can move away for this stalemate. In other words: By adopting a methodology that is more in line with Wittgenstein's later work, we can work towards a better grasp and more detailed description of deep disagreements and their respective hinge commitments.

Accordingly, within my proposed PWEDD, existing WTDD (and their respective HE theories) are to be interpreted as complementarity local descriptions of particular kinds

of deep disagreements (and their respective hinge commitments) in different fields. The choice of one local description rather than another in a given case responds to the particular epistemological features of that deep disagreement and its hinge commitments. Finally, my proposed PWEDD also questions whether the hinge-based and OC focused account of deep disagreements that is assumed by the standard approach of WTDD is the best Wittgensteinian description for deep disagreements in certain fields. For instance, Lavorerio (Forthcoming) has recently shown that deep religious disagreements are better understood by examining Wittgenstein's remarks in the Lectures on Religious Beliefs, where he suggests that these disagreements are a consequence of parties' using different pictures. Along these lines, I examine Wittgenstein's remarks on deep moral disagreements and deep aesthetical disagreements and argue that an adequate description of these disagreements does not benefit from a hinge-based approach.

On the one hand, deep moral disagreements are predicated around different pro/con-attitudes and feelings of approval/disapproval towards certain actions as a consequence of parties' having different ethical systems (Wittgenstein, Rhees & Citron 2015: 28-30; LA: Part 1 §5-7, §35; Rhees 1965: 22-26; PPO: 169/ MT 4.2.[37];). On the other hand, deep aesthetical disagreements are predicated around different pro/con- attitudes and feelings of approval/disapproval towards certain objects as a consequence of parties' having different cultures and, thereby, using different ideals (LA: Part 1 §5-7, §15, §25-29; MWL, 340-343; AWL, §34-§35).

References (with abbreviations)

- [1] Fogelin, R. (1985 [2005]). The logic of deep disagreements. *Informal Logic*, 7 (1): 3-11.
- [2] Kusch, M. (2016). Wittgenstein on Mathematics and certainties. *International Journal for the Study of Scepticism*, 6 (2-3): 120-142.
- [3] Lavorerio, V. (2018). *Do deep disagreements motivate relativism?* Topoi, Special Issue on Disagreement.
- [4] Lavorerio, V. (Forthcoming). *Lectures on Religious Belief and the epistemology of disagreements*. Wittgenstein Studien.
- [5] Moyal-Sharrock, D. (2004). *Understanding Wittgenstein's on certainty*. London: Palgrave Macmillan.
- [6] Moyal-Sharrock, D. (2016). The animal in epistemology. *International Journal for the Study of Skepticism*, 6: 97-119.
- [7] Pritchard, D. (2009). *Defusing epistemic relativism*. *Synthese*, 166: 397-412.

- [8] Pritchard, D. (2011a). Epistemic relativism, epistemic incommensurability, and Wittgensteinian epistemology. In Hales S. D. (Ed.), *The Blackwell companion to relativism* (pp. 266-285). Malden: Wiley-Blackwell.
- [9] Pritchard, D. (2011b). Wittgenstein on scepticism. In Kuusela O, McGinn M (Eds.), *Oxford handbook of Wittgenstein*. Oxford University Press, Oxford.
- [10] Pritchard, D. (2018a). Disagreements, of beliefs and otherwise. In Casey Johnson (Ed.), *Voicing Dissent: The Ethics and Epistemology of Making Disagreement Public* (pp. 22-39). London: Routledge.
- [11] Pritchard, D. (2018b). *Wittgensteinian Hinge Epistemology and Deep Disagreement*. Topoi, Special Issue on Disagreement.
- [12] Ranalli, C. (2020). Deep Disagreement and Hinge Epistemology. *Synthese*, 197: 4975- 5007.
- [13] Ranalli, C. (2018). *What is Deep Disagreement?*. Topoi, Special Issue on Disagreement. <https://doi.org/10.1007/s11245-018-9600-2>.
- [14] Rhees, R. (1965). Some developments in Wittgenstein's Ethics. *Philosophical Review*, 74 (1): 17-26.
- [15] Stroll, A. (1994). *Moore and Wittgenstein on certainty*. Oxford: Oxford University Press.
- [16] Williams M (1991) *Unnatural doubts: epistemological realism and the basis of scepticism*. Blackwell, Oxford
- [17] Williams, M. (2004). Wittgenstein's refutation of idealism, In Denis McManus (Ed.): *Wittgenstein and Scepticism*. London: Routledge.
- [18] Wittgenstein, L. (1967). *Lectures & Conversations on Aesthetics, Psychology and Religious Belief*. California:University of California Press. [LA]
- [19] Wittgenstein, L. (1969). *On Certainty*. Oxford: Blackwell. [OC]
- [20] Wittgenstein, L. (1979). *Wittgenstein's Lectures, Cambridge 1932-35*. Oxford: Blackwell. [AWL]
- [21] Wittgenstein, L. (2000). *Wittgenstein's Nachlass*, The Bergen Electronic Edition. Oxford: Oxford University Press. (References by manuscript/typescript number.)
- [22] Wittgenstein, L. (2013). *Public and Private Occasions*. Maryland: Rowman & Littlefield Publishers. [PPO]

- [23] Wittgenstein, L. (2016). *Lectures, Cambridge 1930-1933: From the Notes of G.E. Moore*. Cambridge: Cambridge University Press. [MWL]
- [24] Wittgenstein, L., Rhees, R. and Citron, G. (2015). Wittgenstein's Philosophical Conversations with Rush Rhees (1939-50): From the Notes of Rush Rhees. *Mind*, 124 (493): 1-71.
- [25] Wright, C. (2004a). Warrant for nothing (and foundations for free)? *Aristotelian Society Supplementary*, 78 (1): 167–212.
- [26] Wright, C. (2004b). Wittgensteinian certainties. In D. McManus (Ed.), *Wittgenstein and Scepticism* (pp. 22–55). London: Routledge.
- [27] Wright, C. (2014). On epistemic entitlement II: Welfare state epistemology. In D. Dodd & E. Zardini (Eds.), *Scepticism and perceptual justification*. Oxford: Oxford University Press.

Mental Disorders, Values and Expressivism

Víctor Fernández*
Universidad de Granada

Miguel Núñez de Prado Gordillo**
Universidad Autónoma de Madrid

In the last decades, there has been an intense debate in philosophy and psychiatry regarding the nature of psychiatric disorders and the role that values, institutions and conventions play in their identification and diagnosis (Hipólito et al. 2018, Kendler 2016, Varga 2015). The positions concerning the nature of psychiatric anomalies have polarized between naturalist views insisting that the psychiatric disorders are biological-based anomalies whose nature is independent of any human conception or activity (Boorse 2014, Wakefield 2007), and constructivist positions claiming that mental disorders are somehow normative constructs (Pickering 2006, Szasz 1961). However, in the last years, several authors (Graham 2010, Thornton 2014) have defended different “having it both ways” positions that attempt to understand mental disorders as value-laden phenomena without dropping a strong commitment with naturalism.

In this paper, we focus critically on Graham’s (2010) view, according to which mental disorders are impairments that affect basic rational operations that, although are not reducible to neural disorders, are objective and universally characterized. Such a view is committed to a particular understanding of values which implies certain universality and individualism: values like rationality or well-being are idealized and invariable standards that prescribe what an individual should be independently of their goals, capacities or their social context; hence why we will refer to it as *value objectivism*. However, as Varga (2015) has argued, this view of values seems to be highly biased by the individualistic conception of agency promoted by western culture. In addition, it overlooks the importance of the social and communal immersion in the acquisition of values, as well as their cultural variability.

We present two lines of criticism against value objectivism. First, we argue that this conception fails to capture the nature of some disagreements regarding the conceptualization, assessment and treatment of mental disorders. Following in previous work (Perez-Navarro et al. 2019), we argue that disagreements involving ascriptions of irrationality and mental disorders exhibit certain features that cannot be captured when one understands values as universal and individualistic standards. This conception seems to

* vfernandezcastro@gmail.com

** miguel.nunnezdeprado@uam.es

be motivated by an aspiration of substituting facts by other ‘smoking gun’ elements that help us to solve disagreements once and for all. In this sense, value objectivism seems to reduce evaluative judgements to mere descriptions of what falls or not under certain universal values. However, we argue that disagreements on mental health assessment and treatment manifest some central features like persistence or appealing to reasons involving further values or metatheoretic reasons that cannot be explained by value objectivism.

Second, we argue that there are two central features of psychotherapy that are incompatible with objectivism about values. First, the management of therapeutic factors in order to “provide patients with the skill to regulate their responses to their thoughts, emotions, and behavior” (Leder 2019: 45), so that their actions do not cause them or others to suffer, or for them to be able to act in accordance with their self-professed values. Second, the attempt to acknowledge patients when they engage in deviant behavior in a non-reactive way, i.e. ascribing responsibility without blaming or accusing the patient (Brandenburg 2018). In our view, those two aspects are not compatible with value objectivism because they presuppose a particularist conception of values, where values are acquired through social dynamics of reactions and co-reactions (McGeer, 2012), and where agents shape each other’s minds in order to scaffold self-regulatory skills. Further, these dynamics are strongly culture-dependent, so the values and norms that shape each other’s minds exhibit an important degree of variability, but also, a dependence of the particular idiosyncratic capacities of the individuals.

To conclude, we sketch an alternative framework to think about mental disorders. Such an alternative is based on a classic *expressivist maneuver* that consists in shifting the question ‘what is a mental disorder?’ to the question ‘what do we do when we attribute a mental disorder?’ Answering the question helps us to focus on perceiving mental disorder attribution as an evaluative practice, where different stakeholders exhibit different attitudes toward the failure to regulate one’s action and cognition in accordance with certain socially shared norms and values in a way that provokes suffering to the ascriber or other agents. This expressivist view incorporates elements of the ‘having it both ways’ views while resisting value objectivism (specifically, the idea that evaluating is just describing what falls under or out of a particular value). Instead, when one evaluates a series of actions as associated with a mental disorder, one is manifesting attitudes and commitments toward such a set of actions. Thus, psychiatric professionals cannot ascribe mental states (nor failures in such mental states) by circumventing their own values and norms, since these cannot be further abstracted from their particular social and individual context. The expressivist view does not however imply some kind of radical relativism according to which we have no way of deciding on specific value-laden assessment matters related to mental disorder attribution and treatment, but it does lead us to abandon any hope of universalism about values, and to encourage another way of thinking about the strategies needed to critically assess and decide what is the best possible decision in each given case and context.

References

- [1] Boorse, C. (2014). A second rebuttal on health. *Journal of Medicine and Philosophy*, 39(6), 683-724.
- [2] Brandenburg, D. (2018). The nurturing stance: Making sense of responsibility without blame. *Pacific Philosophical Quarterly*, 99, 5-22.
- [3] Graham, G. (2013) *The Disordered Mind: An Introduction to Philosophy of Mind and Mental Illness*. Routledge.
- [4] Hipólito I., Gonçalves J., Pereira J.G. (2018) Schizophrenia, Social Practices and Cultural Values: A Conceptual Introduction. In Hipólito I., Gonçalves J., Pereira J. (eds) *Schizophrenia and Common Sense. Studies in Brain and Mind*, vol 12. Springer.
- [5] Kendler, K. S. (2016). The nature of psychiatric disorders. *World Psychiatry*, 15(1), 5- 12.
- [6] Leder, G. (2019). What does it mean to have a meaning problem? Meaning, skill, and the mechanisms of change in psychotherapy. *Philosophy, Psychiatry, & Psychology*, 26(3), 35-50.
- [7] McGeer, V. (2012). Co-reactive attitudes and the making of moral community. In MacKenzie & Langdon (Eds.) *Emotions, Imagination and Moral Reasoning* (pp. 299- 326). Psychology Press.
- [8] Pérez-Navarro, E., Fernández Castro, V., González de Prado Salas, J., & Heras– Escribano, M. (2019). Not expressivist enough: Normative disagreement about belief attribution. *Res Philosophica*.
- [9] Pickering, N. (2006). *The metaphor of mental illness*. Oxford University Press.
- [10] Szasz, T. S. (1961). *The Myth of Mental Illness: Foundations of a Theory of Personal Conduct*. Hoeber-Harper.
- [11] Thornton, T. (2014). Values based practice and authoritarianism. In M. Loughlin (ed.) *Debates in Values-based Practice: arguments for and against* (pp. 50-61). Cambridge University Press.
- [12] Varga, S. (2015). *Naturalism, interpretation, and mental disorder*. Oxford University Press.
- [13] Wakefield, J. C. (2007). The concept of mental disorder: diagnostic implications of the harmful dysfunction analysis. *World Psychiatry*, 6(3), 149-156.

Información y conciencia

Javier Galadí García*
Universidad de Sevilla

José Antonio Langa Rosado**
Universidad de Sevilla

Fernando Soler-Toscano***
Universidad de Sevilla

Presentamos en este trabajo algunos resultados del uso de sistemas dinámicos en el ámbito de los modelos globales de la actividad cerebral para caracterizar estados de conciencia. Nos preguntamos si estos sistemas nos pueden aportar alguna luz acerca de la naturaleza y propiedades de la conciencia y la correspondencia entre las estructuras formales que manejamos y la experiencia consciente. Esto supone cuestionar qué tipo de conocimiento podemos obtener acerca de la actividad cerebral (y de los estados mentales asociados) mediante modelos formales basados en sistemas dinámicos.

Entendemos por conciencia la capacidad de un sujeto para tener una experiencia, con propiedades como información, integración, exclusión, etc. (Tononi y Koch, 2015). Algunas de estas propiedades son análogas a características que se observan en los sistemas dinámicos como la metaestabilidad.

Estructuras y campos informacionales

Consideremos un sistema dinámico como el presentado en la figura 1. Aparece modelado como un grafo de tres nodos interconectados (arriba a la izquierda). Utilizamos estos sistemas para modelar las relaciones entre diferentes áreas del cerebro. La actividad de cada nodo se representa mediante una variable cuya evolución en el tiempo está determinada por una ecuación diferencial. En la figura 1¹ se muestran las ecuaciones de un sistema Lotka- Volterra cooperativo donde la evolución de cada nodo se determina por un parámetro que depende solo del propio nodo (alfa) y la fuerza de las conexiones con los demás nodos con los que mantiene relaciones de cooperación (gamma). En la figura aparece un ejemplo de coeficientes alfa y gamma. En el caso de la actividad cerebral, si cada nodo se corresponde con una zona del cerebro, los coeficientes gamma representan

* javiergaladigarcia@yahoo.es

** langa@us.es

*** fsoler@us.es

¹ Figuras al final en páginas separadas para facilitar la edición.

la conectividad estructural que se puede determinar individualmente. Dados unos parámetros alfa y gamma, el comportamiento del sistema queda condicionado por su atractor, lo que en la figura llamamos Estructura Informacional (ver Esteban *et al.* 2018 y Kalita *et al.* 2019 para más detalles). Contiene todos los puntos estables en la dinámica del sistema y las relaciones de atracción entre ellas.

La Estructura Informacional funciona de esqueleto de una estructura más compleja que es el Campo Informacional (figura 2). En el ejemplo, vemos el que corresponde al atractor de un sistema de dos nodos. Las flechas de la imagen de la izquierda indican la dirección de atracción de cada punto en el espacio. Cuando se asocia a cada punto (del plano en este caso) una función de Lyapunov (niveles de energía) observamos campos informacionales como la imagen de la derecha.

Es posible observar la evolución en el tiempo de las estructuras informacionales en lo que llamamos Estructura Informacional Dinámica (Fig. 3). En un sistema con una conectividad fija (parámetros gamma, tal como ocurre en el cerebro considerando pequeños intervalos de tiempo) el cambio en los parámetros alfa produce distintas estructuras informacionales que van determinando en cada instante la evolución del sistema.

Información y conciencia

Es posible aplicar las ideas anteriores al estudio de la actividad cerebral asociada a distintos estados de conciencia. Hemos usado lo que conocemos como Transformada Lotka- Volterra (Galadí *et al.* 2019) sobre datos de actividad cerebral en pacientes sanos y con alteraciones de conciencia a partir de registros de resonancia magnética funcional en estado de reposo y utilizando para cada uno de ellos su conectividad anatómica (matriz gamma) considerando las siete regiones de la parcelación de Yeo (2011).

En la figura 4 se muestran algunas de las medidas que hemos considerado en las estructuras informacionales que se obtienen en cada instante para cada paciente. Hemos estudiando la distribución de estas medidas en distintos estados (vigilia, sueño) o alteraciones de conciencia (coma, estado de mínima conciencia, etc.), observando que un mayor nivel de conciencia se caracteriza por menor frondosidad (estructuras con menos nodos) y mayores relaciones de cooperación entre las distintas áreas cerebrales.

Conclusiones

Observamos que la actividad consciente se puede asociar con estructuras informacionales con unas características propias. Podemos preguntarnos acerca de la realidad de estas estructuras y su correspondencia con la actividad cerebral, dado que nuestra metodología incluye numerosas simplificaciones como el uso de la parcelación de Yeo o el modelo de Lotka-Volterra cooperativo (que sin ninguna duda es insuficiente para modelar la actividad cerebral). Sin embargo, el teorema fundamental de los sistemas dinámicos

(Norton, 1995) garantiza que cualquier sistema dinámico (más allá del Lotka-Volterra cooperativo) posee un atractor con una estructura que podemos llamar Estructura Informativa. Además, exploramos en qué medida la propuesta de Stanciu (2021) basada en el Realismo Estructural Epistémico como marco para el estudio de la conciencia puede ser aplicable en nuestros modelos.

Referencias

- [1] F.J. Esteban, J. Galadí, J.A. Langa, J.R. Portillo, F. Soler-Toscano, “*Informational structures: A dynamical system approach for integrated information*”, *PLoS Computational Biology* 14(9), 2018: e1006154.
- [2] J. A. Galadí, S. Silva Pereira, Y. S. Perl, M.L. Kringelbach, I. Gayte, H. Laufs, E. Tagliazucchi, J. A. Langa, G. Deco. “Dynamical informational structures characterize the different human brain states of wakefulness and deep sleep”. *bioRxiv*, 2019.
- [3] P. Kalita, J.A. Langa, F. Soler-Toscano, “Informational Structures and Informational Fields as a Prototype for the Description of Postulates of the Integrated Information Theory”, *Entropy* 2019, 21(5), 493.
- [4] D. E. Norton, “The fundamental theorem of dynamical systems”. *Commentationes Mathematicae Universitatis Carolinae*, 36 (3), 1995, 585-597.
- [5] D. Stanciu, An ESR Framework for the Study of Consciousness. *Entropy* 2021, 23, 97. <https://doi.org/10.3390/e23010097>
- [6] B. T. Thomas Yeo, Fenna M. Krienen, Jorge Sepulcre, Mert R. Sabuncu, Danial Lashkari, Marisa Hollinshead, Joshua L. Roffman, Jordan W. Smoller, Lilla Zillei, Jonathan R. Polimeni, Bruce Fischl, Hesheng Liu, Randy L. Buckner. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, 106(3):1125–1165, 2011.
- [7] G. Tononi, C. Koch, “Consciousness: here, there and everywhere?”. *Phil. Trans. R. Soc.*, 2015, B 370: 20140167.

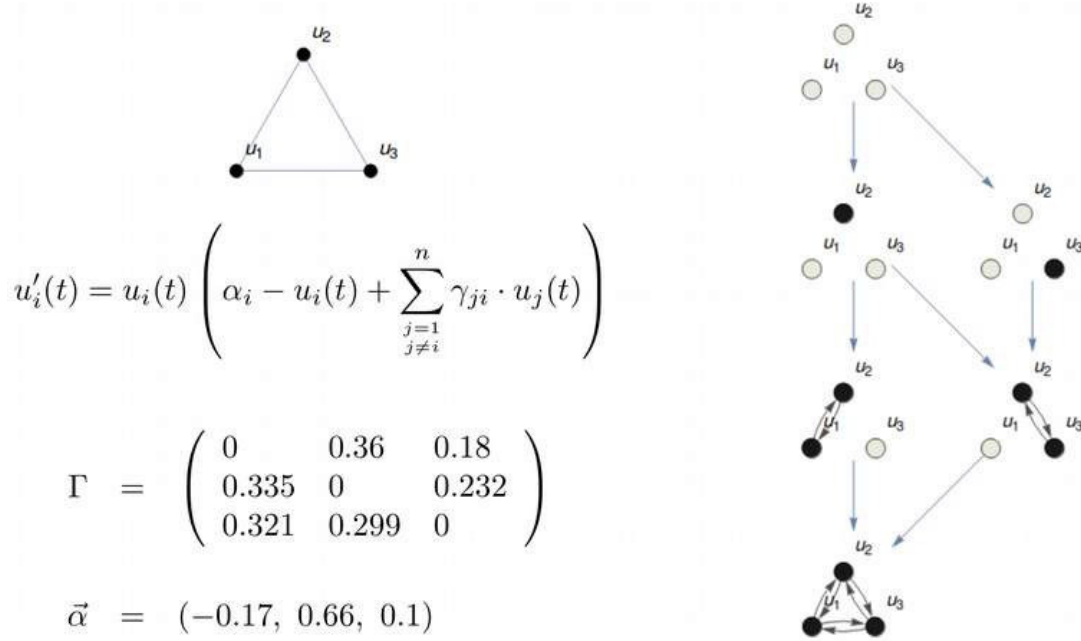


Fig. 1: Estructura informacional de un sistema Lotka-Volterra cooperativo. Arriba a la izquierda se muestra un grafo de tres nodos que modela un sistema cooperativo. El comportamiento de cada nodo viene dado por una ecuación diferencial (izquierda bajo el grafo). Se muestra un ejemplo de matriz de cooperación (parámetros que indican cuánto coopera cada nodo con los demás) y el vector de parámetros alfa (fuerza de cada nodo). A la derecha se muestra la Estructura Informacional que contiene todos los puntos estables de la dinámica del sistema y las soluciones que los unen.

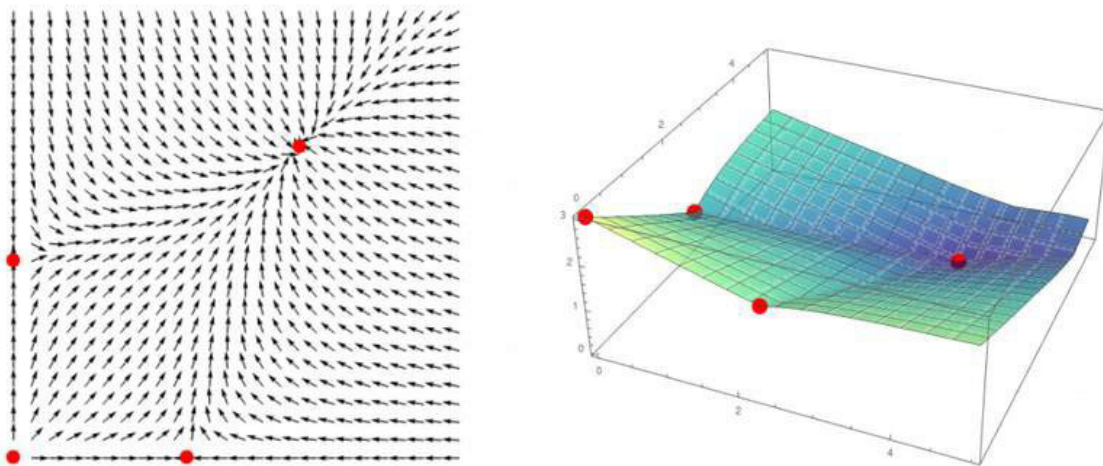


Fig. 2: Campo informacional. Los puntos de la Estructura Informacional dotan de informacióna todo el espacio de estados del sistema. A la izquierda se muestran las direcciones de atracción de cada punto dadas por la estructura informacional de un sistema de dos nodos. Ala derecha se representa gráficamente el campo informacional dado por una función de Lyapunov que muestra los niveles de energía de cada punto del espacio, siendo el punto globalmente estable el que tiene un nivel más bajo.

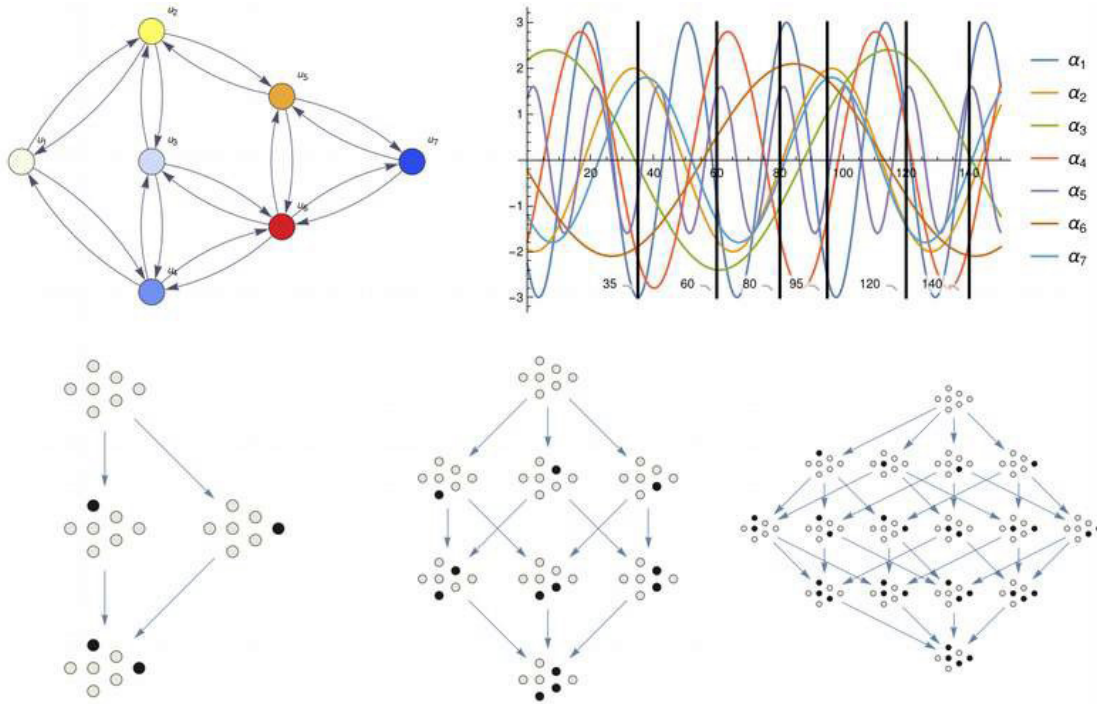


Fig. 3: Estructura Informacional dinámica. Como ejemplo se muestra un sistema de siete nodos (arriba a la izquierda) con una matriz de cooperación constante pero parámetros alfa variables (arriba a la derecha puede verse la evolución a lo largo del tiempo). En función de los parámetros alfa, la evolución del sistema viene dada por una Estructura Informacional diferente (se muestran abajo algunos ejemplos).

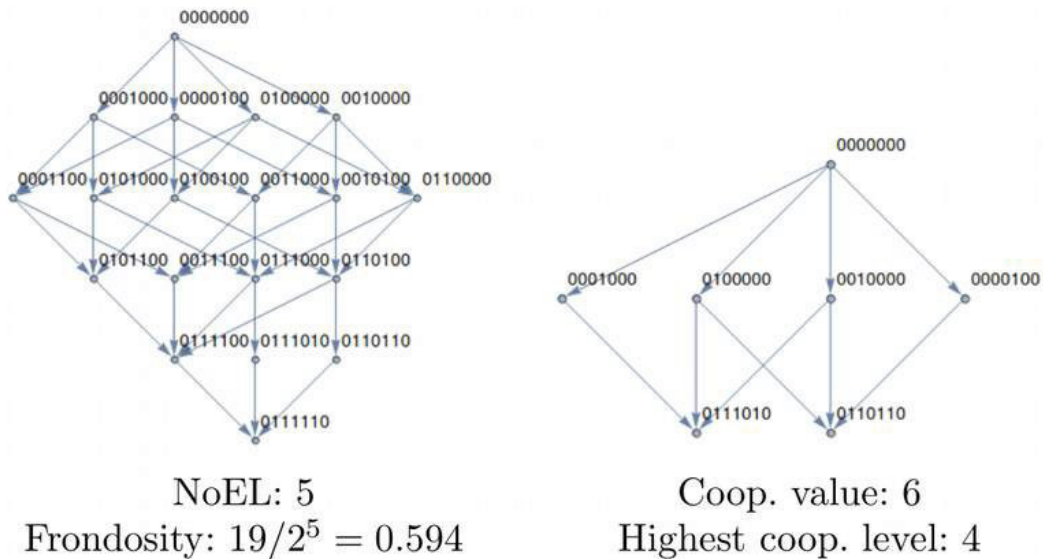


Fig. 4: Medidas sobre estructuras informacionales. A la izquierda aparece una EI correspondiente a un sistema de siete todos y el número de niveles de energía (NoEL) así como la frondosidad. A la derecha se muestran únicamente los nodos de la misma EI en que se dan interacciones cooperativas y se indican los valores de cooperación y el máximo nivel al que se produce dicha cooperación.

Nuevos desarrollos en torno a la posibilidad de una Epistemología feminista de los goznes

Ignacio Gómez-Ledo*
Universidad de Sevilla

En su reciente artículo, Ashton (2019) esboza la posibilidad de llevar a cabo una Epistemología feminista de los goznes. La autora realiza una exploración de posibles vías de encuentro entre ambas corrientes, concluyendo que dicha posibilidad no solo es posible sino además deseable. Sin embargo, como ella misma reconoce, no ha explicado cómo sería en detalle ni los beneficios que dicha interpretación tendría para las Epistemologías feministas, limitando su propuesta a la creación de este nuevo espacio para la reflexión y haciendo un llamamiento a nuevos desarrollos y perspectivas que profundicen en esta línea de investigación. La presente comunicación pretende constituirse como una respuesta a ese llamamiento.

La exposición se desarrollará en el siguiente orden:

- 1) En el primer apartado, analizaremos críticamente los argumentos que ofrece Ashton a favor de una interpretación feminista de la Epistemología de los goznes. Su argumentación se articula en torno a dos desarrollos:
 - Una argumentación que versa sobre la afinidad teórica entre ambas corrientes. Tanto la Teoría del Punto de Vista como la Epistemología de los goznes cuentan con elementos relativos a la justificación epistémica que no responden a la evidencia en el sentido tradicional: el tipo particular de proposiciones llamadas “certezas” en el caso de la Epistemología de los goznes y, en el caso de la Teoría del Punto de Vista, las experiencias diferenciales de los grupos sociales que los llevan a desarrollar recursos epistémicos privilegiados. Además, señala la autora, estos elementos no susceptibles de justificación evidencial presentes en ambas perspectivas son legítimos. Mientras que las corrientes principales en Epistemología tienden a calificar la influencia de factores sociales o no evidenciales en el conocimiento como problemáticos, la Teoría del Punto de Vista los entiende como necesarios y ventajosos aspectos de la justificación.
 - Una exploración de las posibles ventajas de las que puede beneficiarse la Epistemología de los goznes al orientarse hacia los feminismos. Una interpretación feminista de la Epistemología de los goznes deja de estar focalizada en constituirse como una respuesta contra el escepticismo. La justificación

* igngomled@hotmail.com

epistémica es un campo mucho más amplio que el escepticismo y la Epistemología de los goznes cuenta con herramientas conceptuales que pueden aplicarse más allá de este ámbito de discusión limitado.

La Epistemología de los goznes no toma en consideración la literatura y los argumentos sobre el efecto que los factores pragmáticos pueden tener en el dominio epistémico. Explorar la posibilidad de una epistemología feminista de los goznes con una noción de conocimiento socialmente situado puede ser una manera de corregir este error.

La Epistemología de los goznes ha fallado al cuestionar la asunción hegemónica de que el relativismo constituye una objeción de la que escapar en lugar de una posibilidad legítima que desarrollar. Este problema está también en las Epistemologías Feministas, pero algunas propuestas de la Teoría del Punto de Vista han ofrecido tratamientos más creativos del relativismo, lo que podría ayudar a la Epistemología de los goznes a reconsiderar su relación problemática con el relativismo.

- 2) En el segundo apartado ofreceremos nuevos argumentos a favor de la interpretación feminista de la Epistemología de los goznes, reincidiendo en la afinidad teórica entre ambas corrientes, pero poniendo el énfasis en el carácter complementario de sus respectivas características principales.

Baghramian y Coliva (2020) afirman que, a pesar de las diferencias internas entre las diversas corrientes que conforman las Epistemologías feministas, existe una serie de asunciones en común, a saber:

- 1) La carga valorativa de las afirmaciones sobre el conocimiento. La ciencia no es neutral o libre de valores, y estos valores tienen un impacto en la conducta de la ciencia, por lo que deben ser identificados y su influencia reconocida.
- 2) Dependencia contextual del conocimiento. Todas las afirmaciones sobre el conocimiento y su justificación, al menos en parte, son producto de contextos sociales.
- 3) Anti-individualismo. El conocimiento es un producto de comunidades epistémicas antes que de individuos. La justificación se alcanza a través del trabajo epistémico colaborativo de las comunidades.
- 4) Compromiso con una pluralidad aceptable de afirmaciones sobre el conocimiento que pueden entrar en conflicto entre sí.
- 5) Énfasis en el impacto frecuentemente inadvertido del género en las prácticas epistémicas, incluyendo las preguntas que los investigadores priorizan y los tipos de preguntas que valoran.

A estas características habría que sumar una más, que constituye el objetivo fundamental que motiva toda Epistemología feminista: la identificación y cuestionamiento de los mecanismos y estructuras que fundamentan y reproducen la desigualdad de los géneros en el ámbito del conocimiento.

En lo que respecta a la Epistemología de los goznes, no existe un consenso acerca de las características definitorias de la corriente. Debido a que la inspiración de ésta se encuentra en los últimos escritos de L. Wittgenstein (especialmente, *Investigaciones filosóficas* y *Sobre la certeza*) y, dada la manifiesta ambigüedad con la que el autor se expresa en estas obras con respecto a conceptos fundamentales como los de “certeza” o “forma(s) de vida”, existe una pluralidad de interpretaciones de la Epistemología de los goznes (Coliva, 2016).

No obstante, como trataremos de defender, podemos ofrecer una interpretación de la Epistemología de los goznes cuyos rasgos definitorios compartan afinidad teórica con la mayoría de los presupuestos de las Epistemologías feministas (puntos 2, 3 y 4, que he mencionado en la caracterización de Baghranian y Coliva):

- 1) Existe un tipo particular de proposiciones que están fuera de duda (Wittgenstein, 2006: §341).
- 2) Las proposiciones empíricas de cualquier juego de lenguaje dependen de estas proposiciones (Wittgenstein, 2017: §217; Wittgenstein, 2006: §88, §105). Satisface: Dependencia contextual.
- 3) Existe una pluralidad de juegos de lenguaje (Wittgenstein, 2017: §23). Satisface: Pluralidad.
- 4) Las certezas no son individuales, sino que presuponen una comunidad (Wittgenstein, 2006: §298). Satisface: Anti-individualismo.

Si bien es cierto que en la filosofía de los últimos escritos de Wittgenstein no encontramos una reflexión sobre el papel crucial de los valores y el género en el ámbito del conocimiento, no parece que este tipo de asunciones entre en contradicción con cualquiera de los presupuestos fundamentales de la Epistemología de los goznes, por lo que entendemos que la compatibilidad entre ambas perspectivas es posible. Además, tenemos buenas razones para pensar que no solo no hay contradicción, sino que además resulta pertinente investigar el papel de los valores y el género en el ámbito del conocimiento partiendo de la filosofía de Wittgenstein. De hecho, ya existen numerosas investigaciones que tratan de arrojar luz acerca de estos problemas en la actualidad (Tan, 2016; Scheman y O'Connor, 2002).

- 3) Por último, en el tercer apartado propondremos una forma en la que la interpretación feminista de la Epistemología de los goznes puede desarrollarse.

Una vez elucidada la afinidad teórica que habilita la compatibilidad entre ambas corrientes, de lo que se trata ahora es de explorar en qué sentido la Epistemología de los goznes puede aplicarse orientada al objetivo fundamental de las Epistemologías feministas, a saber: identificar y cuestionar las estructuras y mecanismos del poder que mantienen y/o reproducen las desigualdades entre los géneros en el ámbito del conocimiento.

Para ello, nuestra propuesta parte del reconocimiento del papel crucial que juegan las certezas a la hora de constituir y regular nuestras prácticas epistémicas: estas certezas describen nuestra imagen del mundo como trasfondo que me viene dado y sobre lo que distingo lo verdadero de lo falso (Wittgenstein, 2006: §94), y su función es semejante a las de las reglas del juego (Wittgenstein, 2006: §95).

Teniendo esto en cuenta, el objetivo principal de nuestra interpretación feminista de la Epistemología de los goznes sería la identificación y el cuestionamiento de aquellas certezas relacionadas con el género que están detrás de nuestras prácticas epistémicas ordinarias. La posibilidad de cambio en el terreno de las certezas también es algo que Wittgenstein contempla (Wittgenstein, 2006: §96), por lo que el objetivo de cuestionar las certezas que supongan obstáculos a la igualdad y su posterior transformación en certezas que habiliten prácticas epistémicas igualitarias no solo es coherente con los planteamientos teóricos, sino que además supone una aplicación práctica legítima e interesante debido a las nuevas perspectivas que ofrece para los feminismos.

Aunque ya existen intentos rigurosos de clarificar el estatus epistémico de las certezas y sus diferentes tipos (Kusch, 2018), la tarea de identificar aquellas certezas relacionadas con el género es todavía un campo en desarrollo que requiere una investigación en profundidad.

Referencias

- [1] Ashton, Natalie A. (2019): “The Case for a Feminist Hinge Epistemology”, in *Wittgenstein-Studien*, vol. 10, nº 1, pp. 153-163.
- [2] Baghramian, Maria y Coliva, Annalisa (2020): *Relativism*. Londres/Nueva York: Routledge.
- [3] Coliva, Annalisa (2016): “Which Hinge Epistemology?”, en Annalisa Coliva y Danièle Moyal-Sharrock (eds.), *Hinge Epistemology*, Leiden, Brill, pp. 79-96.
- [4] Kusch, Martin (2018): “Disagreement, Certainties, Relativism”, en *Topoi*, pp. 1-9.
- [5] Scheman, Naomi y O’Connor, Peg (eds.) (2002): *Feminist Interpretations of Ludwig Wittgenstein*. Pennsylvania: The Pennsylvania State University Press.
- [6] Tan, Charlene (2016): “Investigator Bias and Theory-Ladenness in Cross-Cultural Research: Insights from Wittgenstein”, en *Current Issues in Comparative Education*, vol. 18, nº 1, pp. 84-95.
- [7] Wittgenstein, Ludwig ([1969] 2006): *Sobre la certeza*. Barcelona: Gedisa.
- [8] — ([1953] 2017): *Investigaciones filosóficas*. Madrid: Editorial Trotta.

Good players in epistemic games

Javier González de Prado Salas*
UNED

What is the relation between norms of inquiry and epistemic norms? Can there be conflicts between these two types of norms, as Friedman (2020) argues? I suggest analyzing epistemic practices by analogy to rule-governed games. Narrow epistemic norms (e.g. evidentialist norms) would constitute standards of correctness for the adoption of attitudes. In turn, inquiry norms would derive from the internal aims of epistemic games (e.g. acquiring knowledge about relevant topics). Crucially, moves that comply with the correctness standards of a game can be inadvisable from the perspectives of the internal aims of that game. In this way, adopting an attitude that satisfies the correctness standards of epistemic practices can be a bad move given the aims of those practices (e.g. it may detract from getting to know interesting truths).

Recent interest in norms of inquiry (e.g. Friedman 2020; Kelp 2021) invites extending epistemic normativity beyond a narrow conception that just includes norms of correctness for belief, or more generally evidentialist norms indicating what attitudes are allowed in response to the evidence (call these ‘narrow epistemic norms’).

However, the extension of epistemic normativity to norms of inquiry is not straightforward. There is an apparent contrast between norms of inquiry and narrow epistemic norms. The latter seem, at least to a large extent, insensitive to practical considerations, in that they cannot be easily balanced against non-epistemic (i.e. wrong kind of) reasons. So, when deliberating one cannot rationally form a belief against one’s evidence just because there is a reward or some other incentive to do so. By contrast, norms of inquiry interact with practical considerations in ordinary ways. Epistemic goals of inquiry can be weighed against other practical aims, so that the reasons derived from norms of inquiry can be outweighed by contrary practical reasons. For instance, one can rationally stop collecting evidence, or inquiring about some topic because of practical reasons to do so (say, a reward, or the risk of harmful outcomes).

This contrast may make us doubt that norms of inquiry are genuinely epistemic. Perhaps the domain of the purely epistemic should be restricted after all to those norms that do not interact with practical considerations. Yet norms of inquiry are linked to assessments of traits, competences, and habits that deserve being called epistemic, insofar as they are essentially related to the attainment of distinctive epistemic goals such as

* jgonzalezdeprado@gmail.com

knowledge or understanding – e.g. open-mindedness, curiosity, impartiality. A view of the epistemic that focuses exclusively on strict evidentialist norms seems unduly restrictive and impoverished (Friedman 2020).

In what sense, then, norms of inquiry can be counted as epistemic? Why do they interact with practical considerations in different ways than evidentialist norms? Are there tensions between inquiry and epistemic norms, as Friedman (2020) contends? In order to answer these questions, I suggest turning to a picture of epistemic practices as rule-governed games (see Maguire and Woods 2020). The idea is that epistemic practices are constitutively governed by norms that determine what attitudes may be correctly adopted in different situations. Narrow epistemic norms are, therefore, constitutive standards of correctness for the relevant attitudes. They would be analogous to the rules determining what moves are allowed in a chess game. However, rule-governed games are not fully characterized by these types of correctness rules. We also need to specify what is the aim of the game, or what outcomes are valuable within the game – what counts as winning, what is a good strategy and what moves are unadvisable given these aims (see Maitra 2011; Marsili 2018). In the case of chess, the ultimate aim of the game is to checkmate the opponent's king. In epistemic practices we can use the label 'epistemic improvement' to refer to the relevant aim (see Friedman 2020), where this could be a matter of figuring out truths, getting knowledge or acquiring understanding. This internal aim of the game determines what moves constitute a good way of playing, that is what moves are conducive to or show proper concern for the aims of the game.

Inquiry norms would derive from these internal aims of epistemic practices – for instance, they can be seen as strategic or instrumental norms regarding the achievement of such aims (Friedman 2020), or as norms about showing concern for what is valuable in the game (see Sylvan 2018). Complying with the correctness standards of the game only guarantees that the agent's moves are allowed, but not that they are intelligent or advisable with respect to the aims of the game. So, one can play extremely bad despite respecting all the correctness standards of the game. For instance, agents that refuse to gather any evidence and just suspend judgment about every relevant issue may be respecting the correctness norms for belief, but they will be poor players of epistemic games, and they will manifest serious flaws in their epistemic competence as inquirers.

The aims of games interact with other goals agents may have while playing the game. So, chess's internal aim to checkmate one's opponent may be balanced against the player's aim to practice certain positions, or to teach their opponent how to play a given opening. Likewise, a football player may pursue the aim of impressing her fans, even over the aim of winning the match. To be sure, these ways of playing may be somewhat defective from the perspective of the internal aims of the game. Yet they can make perfect sense from the viewpoint of agents: they may have good practical reasons to act in such ways.

In general, aims only generate reasons insofar as agents have practical reasons to pursue them. And, of course, these reasons may interact with other reasons the agent has. This is why inquiry norms interact with practical considerations in unconstrained ways. An agent's reasons to inquire about some topic (for instance, to collect relevant evidence) can be outweighed by practical reasons to stop doing so (say, a reward).

Similarly, the correctness standards of games only have normative weight insofar as agents have reasons to respect them – reasons that, following (Maguire and Woods 2020) I take to be ultimately practical. Why is it, then, that in doxastic deliberations epistemic correctness norms cannot be balanced against practical reasons? The answer is that these are constitutive norms of the relevant attitudes, and of the deliberative processes in which they are adopted. Deliberating about what to believe is a matter of figuring out what is the right or correct answer to some question (Hieronymy 2008). So, settling one's doxastic deliberation amounts to treating some answer as right, as correct according to the standards of belief. Therefore, one cannot take oneself to be adopting an attitude like belief while treating this attitude as responsive to practical reasons, in violation of the relevant epistemic correctness standards. In doing so, one would just be adopting a different type of attitude (say, some form of acceptance). Likewise if I treat diagonal moves of rooks as correct, I am not playing the game of chess (I am not playing a game governed by the constitutive correctness norms of chess). By contrast, playing a game perhaps requires that I show some concern for its internal aims, but not that these are my exclusive goals when playing (as pointed out above, I can play a game of chess with the aim of teaching my opponent). Internal aims are not as exclusive and strict as constitutive standards of correctness. Different aims, and the reasons derived from them, can be weighed in a way that standards of correctness cannot (Maguire 2018; Maguire and Woods 2020).

My proposal, therefore, is that narrow epistemic norms are constitutive standards of correctness in epistemic games. In turn, inquiry norms are derived from the internal aims of those games, and tell us what moves are good or advisable in relation to such aims – that is, they tell us how to be good players in epistemic games. I will argue that this proposal provides an attractive account of the apparent tensions between epistemic and inquiry norms discussed by Friedman (2020). Moves that satisfy the correctness standards of epistemic games can nonetheless be bad, inadvisable moves in relation to the internal aims of such games.

Bibliography

- [1] Friedman, J. (2020). The epistemic and the zetetic. *Philosophical Review*, 129(4), 501-536.
- [2] Hieronymi, P. (2008). Responsibility for believing. *Synthese*, 161(3), 357-373.

- [3] Kelp, C. (2021). *Inquiry, knowledge and understanding*. Oxford: Oxford University Press.
- [4] Maguire, B. (2018). There are no reasons for affective attitudes. *Mind*, 127(507), 779-805.
- [5] Maguire, B., & Woods, J. (2020). The game of belief. *Philosophical Review*, 129(2), 211-249.
- [6] Maitra, I. (2011). Assertion, norms, and games. In Brown, J., & Cappelen, H. (eds.), *Assertion: New philosophical essays*. Oxford: Oxford University Press, 277-296.
- [7] Marsili, N. (2018). Truth and assertion: rules versus aims. *Analysis*, 78(4), 638-648.
- [8] Sylvan, K. (2018). Veritism unswamped. *Mind*, 127(506), 381-435.

Reducción y eliminación. Una respuesta a Strawson (2019)

Juan Hermoso Durán*
Centro de Enseñanza Superior Cardenal Cisneros

Las aulas de Wolfson College, en Oxford, acogieron el 25 de mayo de 2017, en el marco de las Isaiah Berlin Lectures, una conferencia del profesor Galen Strawson titulada “One hundred years of consciousness (A long training in absurdity)”. Una versión extendida de aquella conferencia se publicaría en el número correspondiente al primer semestre de 2019 de la revista Estudios de Filosofía. El propósito de este trabajo es discutir algunos de los razonamientos y conclusiones de Strawson, tanto desde el punto de vista de su verosimilitud histórica como desde el de su solidez lógica.

“La Negación” de la existencia de las experiencias conscientes constituye a ojos de Strawson “el episodio más notable de toda la historia de las ideas” (Strawson 2019:10). Dos son las causas inmediatas de este episodio que Strawson trata de aislar, y ambas tienen que ver, a su juicio, con la comprensión defectuosa de una idea por parte de la comunidad filosófica: el conductismo como concepción de la metodología de la investigación psicológica y el naturalismo como concepción de la realidad.

Se apuntará en primer lugar, aunque ello no afecte sustancialmente a su argumento, que la sucinta descripción que Strawson plantea de las relaciones entre el conductismo metodológico y el conductismo filosófico es sumamente débil desde un punto de vista histórico: la tesis según la cual el conductismo de los psicólogos era inmaculadamente metodológico, y su interpretación ontológica un disparate urdido luego por según qué filósofos, no parece, simplemente, ajustarse a los hechos.

En torno de la segunda de las causas de “la Negación” que Strawson anota, un malentendido acerca de las implicaciones del naturalismo, la observación es más bien de índole conceptual. En el transcurso de su razonamiento, Strawson iguala displicentemente reduccionismo y eliminacionismo acerca de lo mental, alegando que reducir lo mental a lo físico es tanto como negar su existencia, dado lo que sabemos de manera transparente sobre la consciencia. El argumento se construye tomando como rival dialéctico a una versión ontológica del conductismo –el “conductismo metafísico enloquecido” de los filósofos, que no sería sino una perversión del saludable conductismo metodológico de los psicólogos, y que aseveraría que “la consciencia no es otra cosa que conducta o disposiciones a la conducta” (Strawson 2019: 15). Tal afirmación, asegura Strawson, equivale

* jhermoso@universidadcisneros.es

a la menos matizada de que “la consciencia no existe”, así que “[...] el conductismo filosófico es de hecho eliminacionista respecto de la consciencia, es decir, niega la existencia de la experiencia –incluso si también niega que lo haga” (íbid.), como suele.

Pero no es cierto, se argumentará, que la diferencia entre reducción y eliminación, que Strawson admite en otros ámbitos del conocimiento científico, decaiga o quede sin efecto en cuanto toca a la naturaleza de la consciencia, ni siquiera dando por evidente todo lo que Strawson considera evidente. En resumidas cuentas: dado lo que sabemos de manera transparente sobre la consciencia, quien postule una tesis reduccionista al respecto –como que la consciencia es un conjunto de conductas o disposiciones a la conducta– se ve comprometido por la lógica de su tesis a probar que las propiedades que sabemos de manera transparente que posee la consciencia quedan explicadas satisfactoriamente por propiedades de aquello a lo que según su tesis se reducen –digamos, conjuntos de conductas o disposiciones a la conducta. Si fracasa en ese empeño, y no logra tampoco convencernos de que lo creemos saber de manera transparente sobre la consciencia en realidad es alguna suerte de ilusión derivada de su verdadera naturaleza, diremos entonces que su tesis es falsa: pero que una tesis reduccionista sea falsa no la convierte en una tesis eliminacionista.

Dicho de otro modo: la distinción entre reducción y eliminación, que Strawson pretende obliterar en el ámbito al menos de la naturaleza de la consciencia, debe mantenerse. Concluir que la consciencia es un fenómeno físico no es lo mismo que negar su existencia. Negar la existencia de la consciencia, o al menos hacerlo de manera fundamentada, requeriría concluir no que la consciencia es un fenómeno físico, sino que ningún fenómeno físico posee las propiedades que atribuimos a la consciencia. Pero eso es precisamente lo que Strawson, acertadamente, niega que sepamos.

¿Por qué un concepto aprendido no puede ser invariante?

José V. Hernández-Conde*
Universidad de Valladolid

1 Introducción

Dos cuestiones principales en el estudio de la naturaleza y estructura de los conceptos son las relativas a su origen y grado de dependencia contextual. Por un lado, en el debate sobre el origen de los conceptos encontramos dos explicaciones principales: (a) *nativista* (o *innatista*), como visión que se extiende desde Platón hasta los nativistas contemporáneos –e.g., Fodor (1981) y Carey (2009)–, según la cual un concepto es innato si está disponible (de modo efectivo, o como disposición¹) desde el nacimiento del sujeto; y (b) *empirista*, como tradición que va desde Aristóteles hasta el empirismo moderno y contemporáneo (Prinz 2002; Shapiro 2004), conforme a la cual un concepto es aprendido si resulta de la acumulación de experiencia sensorial.

Por otro lado, en lo que respecta a su grado de dependencia contextual un concepto puede ser concebido como: (a) *invariante*, si es el mismo para todo contexto e individuo (Keil 1994; Machery 2015); o (b) *dependiente del contexto*, si el concepto considerado no es invariante (es decir, si depende del contexto/sujeto) (Barsalou 1987; Casasanto y Lupyan 2015; Yee y Thompson-Schill 2016).

Las dos cuestiones anteriores (i.e., origen y dependencia contextual de un concepto) suelen ser concebidas como dominios estancos, en el sentido de que adoptar una u otra posición en uno de ellos no condiciona el tipo de postura que puede adoptarse en el otro. Es decir, la asunción de un punto de vista empirista (o nativista) sería compatible con la adopción tanto de un enfoque invariantista, como de un enfoque contextualista, y viceversa.

El objetivo del presente trabajo es mostrar que un concepto no puede ser al mismo tiempo aprendido e invariante, lo que probaría que el origen y el grado de dependencia del contexto de los conceptos no son dominios independientes entre sí.

* jhercon@gmail.com

¹ En el sentido de que el sujeto tenga unos ciertos mecanismos o disposiciones innatos a formar dicho concepto a partir de determinados estímulos adecuados.

2 Planteamiento del problema

Si los conceptos son aprendidos (conforme supone el empirismo), entonces todo concepto es el resultado de una cierta función de aprendizaje, que a partir de un conjunto de entradas produce el concepto en cuestión. Llamemos f a esa función de aprendizaje, I al input experiencial (i.e., conjunto de parámetros de entrada) que conduce a la formación del concepto), y C al concepto producido:

$$f(I) = C$$

Ahora bien, los conceptos producidos son a su vez funciones, capaces de clasificar algo como perteneciente o no a una cierta categoría. Llamemos g a la función asociada a un concepto C que permite clasificar un cierto estímulo E bajo la categoría C (asociada a dicho concepto C) en el seno de un cierto contexto k :

$$C \equiv g(E, k) = \begin{cases} \in C \\ \notin C \end{cases}$$

Obsérvese que tanto I como E son conjuntos de experiencias del sujeto, que solamente han sido distinguidas por ser parámetros de dos clases de funciones diferenciadas (a saber, funciones-de-aprendizaje f , y funciones-concepto g , respectivamente).

3 Definición de las principales posturas

Primero, si un concepto es *innato* (según sostiene el nativismo) entonces estará disponible para el sujeto (de manera efectiva o como disposición) desde su nacimiento. En el primer caso, esto es, si el concepto está efectivamente disponible, eso significa que la función g asociada al concepto C es la misma para todo sujeto s (y momento de tiempo t). En el segundo caso, esto es, si el concepto está disponible como disposición, eso significa que la función f que conduce a la formación del concepto es la misma para todo sujeto s , y conduciría a la formación del mismo concepto $C = g$ si el sujeto ha sido expuesto a algún input experiencial I_i adecuado:

$$f(I_i) = C$$

Segundo, si un concepto es *aprendido* (conforme sostiene el empirismo) entonces existirá una función de aprendizaje f (que sin pérdida de generalidad supondremos la misma para todo sujeto s) capaz de producir una cierta versión i de ese concepto a partir de un cierto input experiencial I_i :

$$f(I_i) = C_i \equiv g_i$$

En este caso cada sujeto s podrá tener una versión i distinta $C_i \equiv g_i$ del concepto, que dependerá de cómo haya sido el input experiencial I_i en que haya basado la formación de dicho concepto.

(Obsérvese que no se puede suponer que el concepto producido $C_i \equiv g_i$ es siempre el mismo, pues eso equivaldría a suponer que el concepto es innato –de modo efectivo o como disposición–.)

* * *

Por otro lado, si el concepto es *invariante*, entonces para todo posible contexto k_u (y sujeto s_v) un cierto conjunto de inputs experienciales E^1 serán siempre categorizados como pertenecientes a C , mientras que otro cierto conjunto de inputs E^2 serán siempre categorizados como no pertenecientes a C . Es decir, para todo contexto k_u (y sujeto s_v) se cumplirá que:

$$g(E_i^1, k_u) = \{\in C, \text{ para todo } E_i^1 \in E^1$$

$$g(E_i^2, k_u) = \{\notin C, \text{ para todo } E_i^2 \in E^2$$

En cambio, si el concepto es *dependiente del contexto*, entonces podrá ocurrir que el mismo input experiencial E sea categorizado en unos contextos como perteneciente a C , y en otros contextos como no perteneciente a C :

$$g(E, k_1) = \{\in C$$

$$g(E, k_2) = \{\notin C$$

4 Empirismo e invariantismo no son compatibles

Considérese ahora el caso de un concepto invariante, y estudiemos cuál es la relación de tal asunción con los dos posibles orígenes de ese concepto.

Comencemos evaluando la posibilidad de que ese concepto fuese *invariante e innato*. En este caso, si el sujeto hubiese estado expuesto a algún input experiencial adecuado I_p , entonces habría producido el concepto $C = g$:

$$f(I_i) = C \equiv g$$

Este concepto categorizaría como pertenecientes a C cierto conjunto de estímulos E^1 , y como no pertenecientes a C otro cierto conjunto de estímulos E^2 , y lo haría así para todo posible contexto k_u (y sujeto s_v):

$$g(E_i^1, k_u) = \{\in C, \text{ para todo } E_i^1 \in E^1$$

$$g(E_i^2, k_u) = \{\notin C, \text{ para todo } E_i^2 \in E^2$$

Por consiguiente, en el caso de un concepto invariante e innato no hay nada nos haga sospechar de que estas dos alternativas pudieran no ser compatibles.

* * *

Consideremos ahora la posibilidad de que el concepto C fuese *invariante y aprendido*. Como antes hemos indicado, no puede suponerse que el concepto producido $C_i \equiv g_i$ es siempre el mismo, pues en ese caso el concepto C sería innato. (El motivo es que si para todo input experiencial adecuado I_i la función de aprendizaje f produce siempre la misma función-concepto $C = g$, entonces el concepto C estaría disponible –al menos como disposición– desde el momento del nacimiento.) Por tanto, existirá un cierto número m de versiones i del concepto, tales que:

$$f(I_i) = C_i \equiv g_i$$

No obstante, siendo el concepto invariante, se cumplirá que para todo posible contexto k_u (y sujeto s_v), y para toda versión $C_i \equiv g_i$ del concepto C :

$$g_i(E_i^1, k_u) = \{ \in C, \text{ para todo } E_i^1 \in E^1$$

$$g_i(E_i^2, k_u) = \{ \notin C, \text{ para todo } E_i^2 \in E^2$$

Ahora bien, el problema estriba en que si dos funciones-concepto g_1 y g_2 producen la misma categorización para cualquier input experiencial E , y para todo posible contexto k_u (y sujeto s_v), entonces estamos realmente ante la misma función g , y lo mismo sucede para las restantes m versiones g_i del concepto. Pero, si el concepto adquirido es siempre el mismo, eso equivaldría a que el concepto considerado no es aprendido, sino innato, en contra de lo que hemos supuesto.

En consecuencia, si un concepto es aprendido no puede ser invariante, y viceversa.

Referencias

- [1] Barsalou, Laurence W. (1987). The instability of the graded structure: implications for the nature of concepts. En U. Neisser (Ed.), *Concepts and Conceptual Development: Ecological and Intellectual Factors in Categorization*. Cambridge: Cambridge University Press, 101-140.
- [2] Carey, Susan (2009). *The Origin of Concepts*. Oxford: Oxford University Press.
- [3] Casasanto, Daniel y Gary Lupyan (2015). All concepts are ad-hoc concepts. En E. Margolis y S. Laurence (Eds.), *The Conceptual Mind: New Directions in the Study of Concepts*. Cambridge, MA: MIT Press, 543-566.

- [4] Fodor, Jerry A. (1981). The present status of the innateness controversy. En J.A. Fodor, *Representations: Philosophical Essays on the Foundations of Cognitive Science*. Cambridge, MA: MIT Press, 257-316.
- [5] Keil, Frank C. (1994). The birth and nurturance of concepts by domains: The origins of concepts of living things. En L.A. Hirschfeld y S.A. Gelman (Eds.), *Mapping the Mind: Domain Specificity in Cognition and Culture*. Cambridge: Cambridge University Press, 234-254.
- [6] Machery, Edouard (2015). By default: Concepts are accessed in a context-independent manner. En E. Margolis y S. Laurence, eds., *The Conceptual Mind: New Directions in the Study of Concepts*. Cambridge, MA: MIT Press, 567-588.
- [7] Prinz, Jesse J. (2002). *Furnishing the Mind*. Cambridge, MA: MIT Press.
- [8] Shapiro, Lawrence A. (2004). *The Mind Incarnate*. Oxford: Oxford University Press.
- [9] Yee, Eiling y Sharon L. Thompson-Schill (2016). Putting concepts into context. *Psychonomic Bulletin & Review* 23: 1015-1027.

Racionalidad y excusas

Julen Ibarrondo Murguialday*
Universidad Autónoma de Madrid

1 Errores, excusas y racionalidad

Los seres humanos no estamos libres de errores, algunos de los cuales podemos excusar apelando a distintos tipos de consideraciones. Una de las formas más habituales de excusarse por haber actuado o creído de manera equivocada consiste en señalar que dicha conducta resultaba racional dadas las circunstancias en las que estaba inmerso el agente. En esta charla, trataré de presentarme brevemente la manera en la que, a mi entender, debemos concebir la función exculpatoria de la racionalidad y desmontar ciertos supuestos bastante extendidos entre los filósofos acerca de en qué consiste esta relación.

Exculpar una acción involucra, en primer lugar, considerarla como errónea. Es cosa distinta el defenderse de una acusación reivindicando lo adecuado de la conducta de uno que conceder que uno estaba equivocado y alegar que su error era excusable (Austin, 1957). Este es sin duda el caso con muchas de nuestras creencias: alguien que olvida algo importante por estar muy distraído con otras preocupaciones, normalmente no afirmará que sus preocupaciones le legitiman a olvidar otras cuestiones importantes. Normalmente, los agobios que alguien pueda tener no convierten en epistémicamente aceptable la pérdida de conocimiento, si bien pueden exculparla.

A menudo comentemos errores a pesar de conducirnos racionalmente. Por racionalidad entiendo aquí, en un sentido amplio, el uso apropiado de nuestras capacidades inferenciales, en virtud de las cuales incorporamos y revisamos estados mentales. Dado que el nivel de racionalidad de una conducta depende del contexto epistémico del sujeto, conducirse racionalmente no destierra la posibilidad de errar. Los agentes pueden carecer de la información relevante y por tanto acabar actuando erróneamente a pesar de conducirse racionalmente. La racionalidad no nos protege ni de la suerte epistémica (Gettier, 1963) ni moral (Williams, 1981). Sin embargo, exhibir racionalidad sí que parece exculparnos de nuestros errores. A fin de cuentas, parece que cuando uno se conduce racionalmente, uno hace cuanto está en su mano para obtener conocimiento y obrar correctamente. Esto parece constituir una muy buena excusa ante los posibles errores (Barceló, 2019).

La idea de que ser racional está íntimamente vinculada a conducirse de la mejor manera posible dadas nuestras limitaciones cognitivas, puede llevarnos a trazar una conexión muy fuerte entre racionalidad y exculpación. Hay dos tesis que trazan este vínculo de manera muy verosímil:

* julen.ibarrondo@uam.es

1. Si es racional que un agente haga ϕ y hacer ϕ es un error, entonces el agente está excusado de este error.
2. Si el agente esté en posición de poder hacer ϕ y hacer ϕ es un error excusable por parte del agente, entonces es racional para el agente hacer ϕ .

Ambas tesis son muy verosímiles y trazan una conexión muy fuerte entre la racionalidad y las excusas. De acuerdo con (1) toda conducta racional es excusable. Y según (2) toda conducta excusable es racional. Sin embargo, como intentaré hacer ver, ambas afirmaciones son falsas.

2 No toda conducta racional es excusable

Muchos autores han defendido que, dado que la racionalidad es una virtud, conducirse racionalmente siempre dice algo bueno acerca del agente (González de Prado Salas, 2018; Sylvan, 2014). Sea esto cierto o no, es obvio que manifestar una virtud no es suficiente para exculpar una conducta: el atracador de bancos sin duda manifiesta virtudes como la valentía o cierto tipo de inteligencia, pero esto no exculpa sus acciones. En realidad, la idea de que la racionalidad siempre exculpa parte de la convicción de que todos los errores de los que somos culpables involucran cierta forma de irracionalidad. Sin embargo, esta idea depende de una concepción excesivamente intelectualista acerca de en qué consiste la culpabilidad de muchos de nuestros errores.

Los juicios estéticos son un buen ejemplo de errores culpables no irracionales. Asumiendo que existen criterios de corrección para los juicios estéticos, parece claro que a menudo manifestamos nuestra falta de gusto de una forma culpable. Aquel que tiene predilección por la decoración hortera y *kitsch* parece cometer un error del que es responsable: hay algo en su mal gusto por el que podemos juzgarlo negativamente. Sin embargo, no está nada claro que dicho mal gusto involucre necesariamente irracionalidad de ningún tipo: su juicio no tiene por qué ser contradictorio ni involucrar creencias desiderativas o debilidad de la voluntad de ningún tipo. Por supuesto hay muchas cosas que ese agente podría hacer para mejorar su gusto y podemos debatir acerca de si debiera haber hecho o no esas cosas. Sin embargo, no parece que hay nada irracional en que no haya dedicado tiempo a formarse en estas cuestiones. A lo mejor podemos excusar su mal gusto alegando que no pudo formarse debidamente en estas cuestiones, pero este tipo de defensa no se escuda ya en lo que es racional para él hacer. A fin de cuentas, pulir el gusto estético no es una obligación racional.

Muchas de nuestras actividades no tienen un componente intelectual muy marcado y el uso de nuestras capacidades racionales no desempeña un papel crucial en ellas: Una persona torpe que pisa a su pareja de baile ciertamente ha cometido un error de apreciación de las distancias y es culpable por ello. Sin embargo, no hay nada particularmente irracional en ser torpe.

3 Tener una excusa no te hace racional

La segunda tesis resulta verosímil porque tendemos a pensar que comportarse racionalmente involucra tatar la información disponible de la manera posible y que esto es suficiente para exculparnos. Asumamos que manejar la información de la mejor manera posible es suficiente para excusarse. Sin embargo, no es cierto que ser racional consista en hacer las cosas de la mejor manera posible. A menudo ciertos factores externos impiden que nos comportemos racionalmente. Si esto es así, conducirse de la mejor manera posible no será lo mismo que ser conducirse racionalmente.

La racionalidad involucra el despliegue efectivo de ciertas competencias (por ejemplo, evitar errores lógicos básicos en el razonamiento). Sin embargo, desplegar una competencia no es lo mismo que actuar lo mejor posible. Imaginemos una persona que, no sabiendo conducir, se ve obligada por una emergencia a tomar el coche. Obviamente, la persona debe intentar conducir de la mejor manera posible (debe de ser cauto, intentar conducir en base a cómo ha visto a otra gente hacerlo, etc.) Sin embargo, es obvio que esta persona no está conduciendo de manera competente: es un pésimo conductor porque no sabe conducir.

Comportarse racionalmente involucra también el despliegue de ciertas capacidades que pueden no estar disponibles en ciertos contextos (Sosa, 2015). Ya sea porque la persona carece de las habilidades correspondientes (como los niños pequeños) o por factores ajenos a su control (drogas, cansancio, etc.). Casos así pueden exculpar nuestros errores, pero eso no convierte a nuestra conducta en racional, de la misma manera que estar exculpadados en nuestra incompetencia en cualquier otra actividad (como el conductor forzado del ejemplo) no nos vuelve competentes.

Conclusiones

Se ve así que la conexión entre excusas y racionalidad es menos directa de lo que normalmente solemos pensar. La idea de que todo error racional es excusable y todo error excusable es racional se basa en una concepción excesivamente intelectualista que exagera el papel de las capacidades racionales en muchas de nuestras actividades, así como de confundir la racionalidad como un logro cognitivo vinculado a ciertas capacidades específicas con el obrar de la mejor manera posible dadas nuestras limitaciones, las cuales pueden ser o no excusables.

En el campo de la epistemología, estas tesis resultan intuitivas dada la tradicional fijación de la disciplina por evaluar la calidad de los juicios en términos de cómo estos se adecuan a la información disponible. Si las tesis (1) y (2) son verdaderas en el campo de la epistemología, solo puede deberse a que hemos restringido la evaluación de la conducta humana a un ámbito en el que las facultades racionales son casi la única habilidad involucrada. No sé si esta forma de concebir las evaluaciones epistémicas es correcta o no, pero

desde luego si las tesis (1) y (2) son verdaderas en el ámbito estrictamente epistemológico no se debe a que sean verdaderas en todos los dominios evaluativos.

Referencias

- [1] Austin, J. (1957). A plea for excuses. *Proceedings of the Aristotelian Society*, 57, 1--30. Barceló, A. (2019). *Falibilidad y normatividad*. Cátedra.
- [2] Gettier, E. L. (1963). Is Justified True Belief Knowledge? *Analysis*, 23(6), 121–123.
- [3] González de Prado Salas, J. (2018). Rationality, Appearances, and Apparent Facts. *Journal of Ethics and Social Philosophy*, 14(2).
- [4] Sosa, E. (2015). *Judgment & Agency*. Oxford University Press UK.
- [5] Sylvan, K. L. (2014). *On the normativity of epistemic rationality*. The State University of New Jersey.
- [6] Williams, B. (1981). Moral Luck. In *Moral Luck: Philosophical Papers 1973-1980*. (pp. 20–39). Cambridge University Press.

Escepticismo, relativismo y puntos de vista: una defensa del perspectivismo alético y meliorista

Andrés Luis Jaume*
Universidad de las Islas Baleares

Tal como aquí pretendo defender podemos ver el el perspectivismo como una vía media entre el relativismo y el absolutismo epistemológicos. Sostiene esta postura epistemológica que nuestro conocimiento se nos da en perspectiva, es decir, que cualquier aserto acerca de la realidad lo que hace es poner de manifiesto una perspectiva del sujeto sobre la misma. El perspectivismo no es una posición novedosa. Ha sido defendido por diversos autores clásicos como Leibniz, Nietzsche u Ortega y Gasset. Recientemente encontramos aproximaciones muy interesantes en el seno de la filosofía analítica como las del propio Hautamäki, A. Moore, E. Sosa o M. Liz y M. Vázquez que se han centrado en la noción de punto de vista.

En líneas generales, el perspectivismo sostiene que nuestro acceso al mundo está mediado por un punto de vista. Pero esta mediación puede entenderse de maneras muy diferentes, en cualquier caso, lo que queda claro es que conocer no es algo así como reflejar un objeto o captar una realidad de una vez por todas, que sería la tesis propia de una concepción absolutista.

El perspectivismo que aquí defiende difiere de alguna de las concepciones contemporáneas precedentes, en particular de la sostenida por Hautamäki (2020) que lo entiende como un tipo de relativismo, así como aquilata el aspecto meliorista destacado por Liz (en elaboración) a la vez que subraya el énfasis puesto en la verdad como objeto del conocimiento y defendido por Sosa (2007; 2009; 2011). Mi objetivo es proponer un perspectivismo comprometido con la idea de verdad que a su vez sea meliorista, es decir, un perspectivismo alético y meliorista . este perspectivismo es capaz de hacer frente al desafío escéptico en su forma relativista a la vez que no es susceptible de posiciones absolutistas o dogmáticas, pues admite que siempre podría ser mejor, siendo su único freno la noción de satisfacción epistémica. Sucintamente, el perspectivismo alético meliorista aquí defendido se compromete con las siguientes tesis:

- (1) Nuestro acceso a la realidad es perspectivista, esto es, no hay una manera absoluta de aprehender los objetos del conocimiento.
- (2) Conocer consiste en realizar un tipo determinado de acción epistémica. La acción epistémica bien evidencia una perspectiva a nivel de conocimiento animal, bien

* andres.jaume@uib.es

elabora una perspectiva a nivel de conocimiento reflexivo apoyada en el conocimiento animal

- (3) En cualquier caso, las diferentes perspectivas que elaboramos sobre los objetos de conocimiento son evaluables y gradativas, es decir, unas son mejores que otras y esto depende, a la postre, de un sentimiento de satisfacción.

La posición aquí defendida debe hacer frente a dos escollos importantes: (1) No debe caer en las redes del escepticismo relativista y (2) no debe caer en la sima de un subjetivismo que nos desconecte de la realidad externa, sino que la noción de satisfacción epistémica debe ser compatible con la idea de un bien.

(1) En primer lugar debe deslindarse del escepticismo y, particularmente del escepticismo relativista. A diferencia del relativismo, el perspectivista no sostiene que todas las perspectivas sean iguales o tengan el mismo valor, respecto del absolutismo, renuncia a llegar a un punto óptimo de verdad, pues siempre piensa que debe seguir investigando y que una perspectiva mejor es posible.

Las relaciones con el escepticismo pirrónico son patentes. Por una parte, el relativismo es un tipo particular de tropo que emplea Sexto Empírico en sus Hipotiposis para suspender el juicio. En efecto, nos dice Sexto que cuando seguimos este tropo “razonamos que, como todas las cosas son con relación a algo, mantendremos en suspensión el juicio sobre el qué son absoluta y objetivamente” (Hip. I, § XIX, 135). Del mismo modo, en el décimo tropo –de la lista de tropos generales para suspender el juicio que consigna en el § XIX– hace referencia a la diversidad de esquemas valorativos y señala que dada la diversidad de formas de vivir y pensar, acceder a tal cosa como la verdad es imposible. Podemos identificar el núcleo de los diferentes relativismo en esos dos tropos de Sexto Empírico. Sin embargo, no por ello debemos rendirnos ante la posibilidad de obtener algún estado de certeza. De otra manera, los tropos de Sexto Empírico no necesariamente tienen que conducirnos a suspender el juicio y renunciar a la búsqueda de la verdad, objetivo final de la empresa epistemológica.

A diferencia del escéptico, quien sostiene alguna forma de perspectivismo no deja de comprometerse con la posibilidad de alcanzar verdades, esto es, con la posibilidad de discriminar alguna relación con el mundo que pueda decirse que es verdadera. El perspectivista, a diferencia del escéptico pirrónico, no vive al margen de la verdad, sino que considera que ésta es una noción muy importante y que vale la pena investigarla, que no es lo mismo estar en la verdad que estar en el error y que, en consecuencia, no es válido el aserto de que hay que conducirse meramente por la observancia de las exigencias vitales –la guía natural, el apremio de las pasiones, el legado de las leyes y costumbres y el aprendizaje de las artes, según nos dice en Hip. I § XI.

Lo que sostengo es que tanto la diversidad de esquemas valorativos como el relativismo son formas bajo las que nos enfrentamos a la realidad y no por ello renunciamos a

hablar de la verdad. Recientemente Hautamäki (2020) ha señalado una serie de precisiones en torno a la noción de verdad. Considera Hautamäki que hay una cierta concepción intuitiva de la verdad y que básicamente ésta concepción se corresponde con la idea de adecuación a la que denomina “verdad común”. De hecho, solemos admitir como verdaderos los informes de las otras personas o nuestros testimonios sensoriales. Es obvio que admitimos la posibilidad de estar equivocados o la divergencia de opiniones, pero todo esto no es óbice para pensar que vivamos en un error permanente. Es más, asumimos la existencia de una realidad compartida que es objeto de nuestros teorías y prácticas sobre el mundo. ¿De qué manera es compatible esta confianza en que es posible el conocimiento, es decir, es posible alcanzar verdades con la idea de un perspectivismo?

La noción de punto de vista ha sido objeto de un análisis detallado (Vázquez y Liz, 2015), se trata de la relación de un sujeto con un objeto a través de cierto aspecto que se considera relevante de la realidad. La realidad es aquello que es independiente del punto de vista, sin embargo, solo aparece como accesible a través de un determinado punto de vista. Este criterio de relevancia puede resultar muy confuso. Alejar la noción de punto de vista de la verdad es hacer que cobre valor la cuestión relativista y, por ende, nos aproxime al escepticismo. Así, debemos asegurar que el resultado de adoptar un punto de vista debe redundar en algún bien epistémico, entendiendo que el bien epistémico supremo es la verdad. Ahora bien, hablar de verdad no debería significar renunciar a nuestra subjetividad. Hay un elemento en relación con la verdad que hace que esta realmente importe: es la satisfacción epistémica, el haber realizado adecuadamente la tarea propuesta.

Cuando ejecutamos una acción epistémica lo que pretendemos es un acierto debido a nuestra competencia como agentes epistémicos (Sosa 2007). Cualquier ejecución epistémica es susceptible de ser evaluada. Uno puede aproximarse más al blanco o menos, dar en el blanco o no dar, lo importante es que el tiro resulte de la competencia del sujeto y la cuestión del acierto no se deje en manos del azar. Tender a la verdad es tender a un bien epistémico y su ejercicio constante hace que devenga hábito. Podemos reconocer los buenos hábitos, las virtudes y diferenciarlos del vicio. Sin embargo, lejos de una concepción excesivamente objetivista, podemos también introducir una variable como es la satisfacción. En efecto, la satisfacción es algo que corresponde al propio sujeto. Uno puede decir que está satisfecho o insatisfecho. La satisfacción tiene un elemento de autonomía que no debería perderse aunque tampoco diluirse en la sima de la subjetividad. Lo que buscamos es una satisfacción epistémica razonable, no caprichosa. El sujeto adopta un punto de vista conforme a una finalidad, la verdad; no una verdad desnuda, sino una verdad que se interesa por un modo de estar en el mundo que considera bueno. Así adoptar un punto de vista supone ver la realidad de algún modo, pero nada impide que podamos considerar otra manera mejor de ver esa misma realidad. El punto en el que nos detenemos es aquel en el que nuestras expectativas resultan razonablemente satisfechas.

Bibliografía

- [1] Hautamäki, A. (2020), *Viewpoint relativism*, Berlin, Springer.
- [2] Liz, M. (2013), “Analizando la noción de puntos de vista” en Liz (Ed.) *Puntos de vista. Una investigación filosófica*, Barcelona: Laertes.
- [3] Liz, M. (en elaboración), “Algo en que creer. Naturalismo, teísmo y perspectivas personales”
- [4] Sosa, E. (2007), *A Virtue Epistemology: Apt Belief and Reflective Knowledge*, vol I, Oxford Univ. Press.
- [5] Sosa, E. (2009), *Reflective Knowledge: Apt Belief and Reflective Knowledge*, vol II, Oxford Univ. Press.
- [6] Sosa, E. (2010), *Knowing Full Well*, Princeton, Princeton Univ. Press.

Verdad relativista y verdad perspectivista

Manuel Liz*
Universidad de La Laguna

No suele considerarse que puede haber una concepción perspectivista de la verdad con peso similar al de otras concepciones. Argumentaremos que esto es un error. Lo haremos analizando dos temas importantes dentro del relativismo alético planteado en los últimos años: la retractación (retraction) y los desacuerdos sin tacha (faultless disagreements). Mostraremos cómo ciertas tesis propias del perspectivismo explicarían de modo muy plausible la existencia de una disposición hacia la retractación. Y respecto a los desacuerdos sin tacha, mostraremos que el perspectivismo dispone de recursos que no estarían al alcance del relativismo.

1 Retracción

Michela Massimi (2016) ha analizado recientemente cuatro opciones a la hora de precisar la noción de verdad perspectivista. En la primera, las perspectivas son constitutivas de los contenidos proposicionales. En la segunda, son constitutivas de sus valores de verdad. En la tercera, son constitutivas de las condiciones para asertar tales proposiciones en los contextos de producción. Y en la cuarta, son constitutivas de esas condiciones de aserción también en relación a otros contextos de evaluación.

Las dos primeras opciones son relativistas extremas. La tercera opción es adoptada por los relativismos contextualistas, históricos, etc. Y también es una forma bastante extrema de relativismo. Massimi argumenta que únicamente la cuarta opción consigue distanciar al perspectivismo de esos relativismos, dando sentido al ideal realista de “getting things right” sin colapsar tampoco en formas extremas de realismo.

Esta última opción permite una continuidad a través de los contextos de las condiciones de aserción de los contenidos proposicionales. Massimi se apoya en el relativismo alético de MacFarlane. En ciertas áreas del discurso, nuestras afirmaciones de conocimiento involucran compromisos que las hacen evaluables desde otras perspectivas diferentes a las perspectivas desde las que de hecho son formuladas.

Con MacFarlane (2014), diríamos que en esas áreas del discurso está activa una disposición a retractarnos de nuestras afirmaciones de conocimiento si se demuestra convincentemente su falsedad. Y esto implica que las afirmaciones de conocimiento son

* manuliz@ull.es

evaluables desde otras perspectivas, rivales o posteriores. (Ver Kölbel 2002, García-Carpintero y Kölbel, eds., 2008, y Marques 2018).

Dentro de la ciencia, tal mecanismo es fundamental. Está en el corazón mismo de la actividad científica. Da sentido a la crítica mútua y al progreso en el conocimiento.

El problema es que si el perspectivismo no tuviera mayor sustantividad, no se distinguiría apenas de un relativismo moderado, circunscrito a ciertas áreas del discurso, como el planteado por MacFarlane. Ni tampoco, de un realismo moderado que asumiera esas mismas tesis.

Esa es la imagen que encontramos en muchos autores partidarios de un perspectivismo científico, como Evandro Agazzi, Sandra Mitchell o Ronald Giere. Y también está presente en la propia Massimi.

Pero puede argumentarse que el perspectivismo implica ciertas tesis que sí le darían sustantividad (Vázquez y Liz, eds., 2015). Podemos considerar, por ejemplo

- T1 Dependencia perspectivista: Todos nuestros accesos a la realidad dependen de nuestros PV en el sentido de que no podemos entender cómo podría existir esa realidad aparte de la manera como se nos presenta desde los PV que adoptamos.
- T2 No igualitarismo: No todos los PV tienen el mismo valor.
- T3 Meliorismo: Siempre es posible adoptar PV mejores, es decir PV con más valor.
- T4 Invariantismo: Podemos reconocer la existencia de invarianzas perspectivistas relevantes, especialmente en los contenidos explícitos de nuestros PV.
- T5 External World: A través de una gran variedad de PV se nos presenta un mundo de objetos y de sujetos teniendo propiedades, manteniendo relaciones y constituyendo hechos, eventos y procesos que no dependen de nosotros, que somos capaces de percibir de diferentes formas, y sobre los cuales podemos pensar y actuar.

T1 es contraria a todo realismo, incluyendo los realismos moderados. Y quedaría reforzada con T3. T2 es contraria a todo relativismo, incluyendo los relativismos moderados. De nuevo, T3 reforzaría ese rechazo. Todo el contenido realista del perspectivismo quedaría condensado en T4 y T5 (un caso particular de T4). Para el perspectivismo, esto no sólo es suficiente sino lo único que cabe obtener (Nozick 2003, Hales 1993).

T3 (muy presente en Leibniz, James y Putnam) tiene un papel muy importante en el perspectivismo. Actúa como motor de la dinámica de perspectivas. Adoptar una perspectiva y cambiar de perspectiva son procesos cuya orientación general es la búsqueda de perspectivas mejores.

Tal meliorismo ofrece una explicación sumamente plausible de porqué puede activarse el mecanismo de retractación. Para el perspectivismo, habría en realidad dos mecanismos, uno de retractación y otro de reafirmación.

El segundo mecanismo promovería la “fidelidad a la propia perspectiva”. Esta idea también ha tenido un gran peso en la tradición perspectivista. En el caso de las creencias, implicaría valorar muy positivamente todo aquello que nos reafirme en lo que creemos.

De acuerdo al perspectivismo, siempre existiría una relación dinámica entre retractación y reafirmación. Y hay contextos normativos e instituciones que hacen que prevalezca la retractación sobre la reafirmación. Uno de esos contextos es la ciencia (otros son el ámbito jurídico y los deportes de competición). Podemos decir que la ciencia se ha ocupado de regular las retractaciones que tienen que ver con el conocimiento.

Destacan dos consecuencias: 1) la prevalencia o equilibrio entre reafirmación y retractación será siempre dinámica y en alguna medida inestable, 2) la distinción entre áreas del discurso donde la retractación juega un papel importante y áreas del discurso donde no lo juega es también una distinción dinámica e inestable.

2 Desacuerdos sin tacha

Hay áreas del discurso en las que nuestras afirmaciones no involucran ninguna disposición a la retractación. Dentro de las opciones ofrecidas por Massimi, volveríamos a la casilla 3.

Los juicios de gusto, los juicios éticos o estéticos, las proposiciones sobre el futuro, etc., serían algunas de esas áreas. En ellas hace su aparición el fenómeno de los desacuerdos sin tacha. Su estructura general es la siguiente:

Un sujeto S1 afirma p. Otro sujeto S2 afirma q. Hay una contradicción entre p y q. Sin embargo, todas las evidencias y razones que pueden esgrimir tanto S1 como S2 parecen impecables.

Estos conflictos sugieren aplicar un concepto de verdad relativa. Relativizar las afirmaciones de S1 y S2 a sus perspectivas disuelve la contradicción (Goodman 1978, Boghossian 2006). Y la posible activación del mecanismo de retractación limitaría el efecto de esa relativización sin recurrir al realismo.

De nuevo la pregunta es si el perspectivismo no puede ofrecer más que esto. Únicamente una consideración muy superficial del perspectivismo conduce a pensar que los casos de desacuerdos sin tacha sólo pueden ser tratados introduciendo las perspectivas como nuevos índices de evaluación y añadiendo algún mecanismo como la retractación.

Otras estrategias perspectivistas serían:

1. Meliorismo creando desigualdades

T3 es una tesis característica del perspectivismo. Pero no desempeña ningún papel en el relativismo. Y algunas de las desigualdades generadas podrían cambiar la situación, afectando a las evidencias o a las razones.

Tratar los desacuerdos sin tacha de una manera más dinámica permitiría tomar seriamente esta posibilidad.

2. Puntos de vista reflexivos

Adoptar un punto de vista reflexivo siempre requiere asumir una posición no comprometida sino distante. Esto se aplica también cuando nuestra reflexión concierne a los puntos de vista que adoptamos. Y cuando nuestra posición es distante, cambian con frecuencia nuestras evaluaciones.

Nuevamente, un tratamiento más dinámico daría relevancia a este papel de la reflexión

3. Puntos de vista en interacción

Cuando dos puntos de vista en interacción tienen contenidos contradictorios, suele darse una divergencia en las actitudes que se mantienen hacia esos contenidos. Y cuando dos puntos de vista contienen las mismas actitudes, suele evitarse la aparición de contenidos que se contradigan.

Esa dinámica es esencial en nuestras prácticas discursivas, especialmente cuando la interacción entre los puntos de vista tiene lugar en contextos dialógicos.

4. Áreas del discurso en construcción

Ya hemos indicado que hay áreas del discurso abonadas para que surjan situaciones de desacuerdo sin tacha. Pero el perspectivismo rechazaría que esas áreas estén consituadas con independencia de nuestras perspectivas. Han de ser vistas tan sólo como invarianzas perspectivistas. Y los casos particulares seguramente puedan tener mucha más inestabilidad.

Pensar que existen evidencias y razones definitivas conduciendo desde distintos puntos a contradicciones también es propio del realismo. Y el perspectivismo lo rechazaría con los mismos argumentos que rechaza todo realismo.

El perspectivismo tiene más recursos conceptuales que el relativismo (Hautamäki 2020, cap. 8). En todos ellos hay una idea estratégica muy importante. Los conflictos de perspectivas se resuelven siempre dinámicamente teniendo en cuenta otras perspectivas.

Referencias

- [1] Boghossian, Paul (2006). *Fear of Knowledge*. Oxford: Oxford Univ. Press.
- [2] García-Carpintero, Manuel, y Max Kölbel (eds.) (2008). *Relative Truth*. Oxford: Oxford Univ. Press.
- [3] Goodman, Nelson. (1978). *Ways of Worldmaking*. Indianapolis: Hackett.
- [4] Hales, Steven, 1997, "A Consistent Relativism", *Mind*, 196, 421: 33-52.
- [5] Hautamäki, Antti (2020). *Viewpoint Relativism*. Heidelberg: Springer.
- [6] Kölbel, Max (2002). *Truth without objectivity*. Londres: Routledge.

- [7] MacFarlane, John (2014). *Assessment Sensitivity*. Oxford: Oxford Univ. Press.
- [8] Marques, Teresa (2018). “Retractions”. *Synthese*, 195: 3335-3359.
- [9] Massimi, Michela (2016). “Four Kinds of Perspectival Truth”. *Philosophy and Phenomenological Research*, XCVI, 2: 342-359.
- [10] Nozick, Robert (2003) *Invariances*. Boston: Harvard Univ. Press.
- [11] Vázquez, Margarita and Manuel Liz (eds.) (2015). *Temporal Points of View*. Heidelberg: Springer.

Playing possum: Animal thanatosis and the concept of death

Susana Monsó*

Messerli Research Institute, University of Veterinary Medicine Vienna

The concept of death has traditionally been viewed as a uniquely human feat. Monsó (2019) has recently argued against this view by reasoning that the concept of death should not be construed in all-or-nothing terms, but ought to be conceived rather as a spectrum. While the concept of death would have varying levels of complexity, there is a bare minimum that an animal needs to grasp in order for us to credit her with a concept of death. This is the notion of irreversible non-functionality: the idea that dead individuals do not do the sorts of things that we expect living beings to do and that this state is permanent. In this paper, I use Monsó's framework to argue that the defence mechanism known as thanatosis provides us with strong evidence of the prevalence of the concept of death, thus understood, among nonhuman animals.

Thanatosis, or death feigning, is a behaviour that has emerged several times throughout evolutionary history. It can be distinguished from another similar behaviour called tonic immobility. While tonic immobility consists primarily of a paralysis, in thanatosis we see an active imitation of the characteristics of corpses. The Virginia opossum, for instance, not only becomes paralysed when threatened, but also imitates the bodily and facial expressions of a corpse, reduces her body temperature and heart rate to a minimum, her tongue turns blue, and her anal glands simulate the smell of rot (Francq, 1969). In snakes of the genus *Heterodon*, thanatosis is preceded by a dramatic 'death' display and can include the secretion of blood from the mouth (Burghardt, 1991).

Thanatosis and tonic immobility are distinguished, not just in terms of their complexity, but also because of the way they function in providing a defence from predators. Predation is a sequence made up of four phases: detection of the prey, recognition of it as suitable for eating, subjugation, and consumption (Rojas and Burdfield-Steel, 2017). Thanatosis and tonic immobility both act in different phases.

Tonic immobility is, firstly, an anti-detection mechanism. This means that it helps the prey not to be perceptually detected by the predator. By becoming completely still, the prey can become one with the environment, and so tonic immobility can aid those animals that engage in camouflage. In addition, many predators need to detect movement in order

* susanamonso@gmail.com

to initiate their predatory behaviour, and so tonic immobility serves to eliminate a stimulus that predators are attracted to.

Tonic immobility also serves as an anti-consumption mechanism. For instance, some toxic frogs assume a bodily posture with their limbs folded against their body that reduces their chances of getting hurt when swallowed and allows for them to be spit out whole (Toledo et al., 2010). Some insects also assume very rigid postures with their limbs extended that makes it difficult for some predators that do not chew to swallow them (Honma et al., 2006).

Thanatosis differs from tonic immobility in that it functions primarily as a mechanism to prevent, not detection and consumption, but recognition and subjugation.

As an anti-recognition mechanism, it works by making the prey seem unappetising. It would work on those predators who have learned in the past that dead prey is unpalatable or upsetting for the stomach. It can also serve to reduce or eliminate the motivation of those predators who specialise in hunting live prey. By classifying the animal as dead, the predator would prefer to find another target.

As an anti-subjugation mechanism, thanatosis would work by relaxing the predator's attention on the prey. Some predators temporarily store their prey before consumption. Others may attempt to kill more than one animal in a short amount of time and there may be a cost to ensuring that each animal is actually dead. All of this can give the prey who is feigning death a chance to escape (Gregory et al., 2007; Humphreys and Ruxton, 2018).

Regardless of whether it is functioning as an anti-recognition or an anti-subjugation mechanism, thanatosis works by taking advantage of predators' concept of death—of her capacity to process that the prey is irreversibly non-functional.

One might object that thanatosis normally involves a paralysis too, and so could work instead as an anti-detection mechanism. While this is a possibility, it remains unlikely that this is the primary reason why thanatosis evolved. This is, firstly, because thanatosis is triggered upon contact with the predator, which presupposes that the prey has already been detected. But more importantly, the very characteristics of complex thanatosis displays point to the predators' concept of death being the selective pressure that gave rise to it.

The defence mechanisms that prey have evolved provide us with an insight into the mental lives of their predators, because these constitute the selective pressure that has shaped those mechanisms. For instance, the existence of insects who are shaped like leaves or sticks shows us that there are predators who are not interested in eating leaves or sticks. Likewise, the Batesian mimicry that we see in some flies, who have evolved black and yellow stripes like those of toxic species like wasps or bees, shows us that there are predators who can form an association between black-and-yellow-stripes and toxicity.

Thanatosis displays are made up of a cluster of heterogeneous behaviours whose only thing in common is that they are characteristics of corpses. The only selective pressure that could have given rise to them is the concept of death of the deceived predators.

That thanatosis is evidence of predators' concept of death is likewise entailed by the fact that it is a defence mechanism that would not work against specialised predators, since they would have developed an adequate response. Instead, it is likely to work against generalist predators who do not encounter such prey too often and are not familiar with their trick. For it to work against this kind of predators, we need them to have not just the capacity to react to certain concrete stimuli that are associated with death, but rather the ability to classify different individuals as possessing the property 'dead'. This is in line with what Monsó (2019) requires for the possession of a concept of death.

Monsó and Osuna-Mascaró (2020) have argued that large predators are strong candidates for the possession of a concept of death. This is because they tend to be cognitively complex, with both a tendency and an incentive to monitor the non-functionality of prey. A prey animal that is alive poses a potential danger, while a dead one ensures a full stomach. The low success rate of predators, coupled with the existence of a myriad of defence mechanisms in prey animals, ensures that they will pay attention to the moment of death. In addition, predators, especially carnivores, have daily opportunities to learn about death.

The distribution of thanatosis in the animal kingdom is a further reason to believe that the concept of death will be prevalent among predators. We find elaborate forms of thanatosis in, at least, amphibians, reptiles, birds, and mammals (Humphreys and Ruxton, 2018). While their presence in the phylogenetic tree is very extended, it is also patchy, which suggests that thanatosis has appeared through convergent evolution in species that are only distantly related. This means that these different species have faced similar selective pressures in their different habitats. This points us to the idea that the concept of death, rather than pertaining solely to the human species, is likely a fairly common cognitive trait in the animal kingdom.

References

- [1] Burghardt, G. M. (1991). "Cognitive ethology and critical anthropomorphism: A snake with two heads and hognose snakes that play dead", in *Cognitive ethology: The minds of other animals*. New Jersey: Lawrence Erlbaum Associates Publishers, pp. 53-90.
- [2] Francq, E. N. (1969). "Behavioral Aspects of Feigned Death in the Opossum *Didelphis marsupialis*", *The American Midland Naturalist*, 81(2), pp. 556-568.
- [3] Gregory, P. T., Isaac, L. A. and Griffiths, R. A. (2007). "Death feigning by grass snakes (*Natrix natrix*) in response to handling by human 'predators'", *Journal of Comparative Psychology*, 121(2), pp. 123-129.

- [4] Honma, A., Oku, S. and Nishida, T. (2006). “Adaptive significance of death feigning posture as a specialized inducible defence against gape-limited predators”, *Proceedings of the Royal Society B: Biological Sciences*, 273(1594), pp. 1631-1636.
- [5] Humphreys, R. K. and Ruxton, G. D. (2018). “A review of thanatosis (death feigning) as an anti-predator behaviour”, *Behavioral Ecology and Sociobiology*, 72(22).
- [6] Monsó, S. (2019). “How to Tell If Animals Can Understand Death”, *Erkenntnis*. doi: 10.1007/s10670-019-00187-2.
- [7] Monsó, S. and Osuna-Mascaró, A. J. (2020). “Death is common, so is understanding it: the concept of death in other species”, *Synthese*. doi: 10.1007/s11229-020-02882-y.
- [8] Rojas, B. and Burdfield-Steel, E. (2017). “Predator Defense“, in Vonk, J. and Shackelford, T. (eds), *Encyclopedia of Animal Cognition and Behavior*. Cham: Springer International Publishing, pp. 1-8. doi: 10.1007/978-3-319-47829-6_708-1.
- [9] Toledo, L. F., Sazima, I. and Haddad, C. F. B. (2010). “Is it all death feigning? Case in anurans“, *Journal of Natural History*, 44(31-32), pp. 1979-1988.

Secrets: An Epistemological Account

Jesús Navarro*
University of Seville

Ignorance has been drawing a great deal of attention recently in both social and political epistemology¹. In contrast, secrecy has gone almost unnoticed in those fields, in spite of being a certainly social and political issue². Perhaps the problem is that it is far from obvious why secret is of *epistemological* concern³. With this paper I intend to pave the way for an epistemological study of secrecy, proposing an account of that phenomenon as a particular kind of disturbance in the social flow of knowledge.

As an epithet, “secret” qualifies all sorts of things (there are secret facilities, secret plots, secret feelings, secret recipes,...). Occasionally, “secret” may also qualify “knowledge”, but not very often. However, In this paper I will present and defend the strongly epistemic account that whenever something is a secret, it is known. *All secret is both known by somebody and prevented to be known by somebody else*, that could be the lemma.

An alternative view would hold that it is not knowledge, but mere information about those things, properties or states what is at stake—a view that I dispel in the first section of this paper. My claim implies, among other things, that the state of being a secret is factive (if one keeps a secret, it must be true). But the thesis is even stronger than this: in my view, secrets have all the conditions required for knowledge: not only factivity, but also justification, and being Gettier-proof. Somebody may not have a secret merely because she *believes* that she has one, or merely because she acts as if she had one: the one who really keeps a secret must actually *have* what she conceals—i.e., she must know it.

A second section is devoted to the presentation of a detailed reductive analysis, which is the core of my account:

* jnr@us.es

¹ Several collective volumes on the topic bear witness, such as Proctor and Schiebinger (2008), Gross and McGoey (2015), Peels and Blaauw (2016), or Barton and Davis (2018). Ignorance, mostly when it is produced or manufactured, has become central in social epistemology thanks gender and race studies, mostly after Mills (1997); see Sullivan and Tuana (2007). In contrast, secret has remained a rather marginal topic in this context, although see Ahmed (2010).

² An exception in this respect is Broncano (2019, 2020), whose employment of the concept motivates the present view.

³ Secrecy has become a hot topic in cultural studies recently (see Birchall 2011, 2014, 2016), up to the point that a monographic journal, *Secrecy Studies*, been devoted to it since 2016. Secrecy is also an issue in political philosophy, following concerns for privacy, transparency and civil rights (see Resnik 2006, Borradori 2016). However, these other contexts are populated with information-based accounts, and what remains obscure in them is the relationship between secrecy and knowledge.

Something is a secret iff it is

- (1) known by someone –the *bearer*–,
- (2) ignored by someone –the *ignoramus*–,
- (3) concealed by someone –the *keeper*– in such a way that
- (4) the epistemic interests of the *ignoramus* are frustrated.

In section 3 I dig into the concept of secret by telling it apart from the one of *mystery*. Mysteries are situations where conditions (2) and (4) are met: somebody ignores something and sees her own epistemic interests frustrated. This property is relational. Propositions are not mysteries in themselves, but to someone. Some proposition *p* maybe a mystery for *A* but not for *B*, either because *A* has worse epistemic faculties, *A* in a worse epistemic position, or simply *A* has some epistemic interests that *B* does not. Being a mystery is a function that relates the agent's epistemic capacities and interests to some putative knowledge. So defined, secrets are a kind of mystery: situations where not only conditions (2) and (4) are met, but also conditions (1) and (3).

The difference between mysteries and secrets is easy to explain to anyone who has played the famed *Cluedo* game. On the one hand, three cards are kept in an envelope in the centre of the table, cards that no one has ever seen. Those cards are, in my terms, a pure and simple mystery to the players. On the other hand, each player is given several cards, which are not just mysteries, but the secrets of the game. At the beginning, each player knows what cards she has and, as the game goes, that knowledge is revealed to some players, but concealed to some others... knowledge flows, but it is intentionally prevented from flowing freely.

In section 4 I show that condition (2) locates the topic of secrecy inside the fruitful field of the epistemology of ignorance, or agnotology, but at the same time it defines its proper locus inside that discipline. Not all cases of ignorance are instances of secrecy. That is why conditions (1) and (2) fall short of defining the category of secrets, which explains the specific lacuna that social and political epistemologies have had until now.

In section 5 I minimally elaborate upon the idea that, in many cases, some of the three characters that I have identified in the situation of secrecy may coincide.

In section 6 I discuss condition (4), which requires the *ignoramus* to have—either actually or potentially—some epistemic interest that is frustrated by the *keeper*'s performance or lack thereof. Again, the putative *keeper* would not be keeping any genuine secret after all if no one had any interest in it. I am aware that this may seem controversial, and I try to dispel this worry by analogy with the problem that pointless knowledge presents for the existence of intrinsic epistemic value.

According to the account that I propose, secret is, by its very nature, both a social and an agential phenomenon. It takes two to have a secret, which is its social component (although I discuss possible exceptions in the paper); and for something to be a secret someone, the *keeper*, must actually *make* it a secret, which is the agential component.

Consequently, the issue of *legitimacy* arises, and I focus on it in section 7. This forces us to understand secrecy in the context of social normativity, where epistemic rights and duties are distributed, and either respected or violated. This section ends with a quick look at the way secrecy fits into some heated topics in social epistemology: epistemic injustice, the proliferation of fake news, and conspiracy theories.

Independently of its legitimacy, there are countless ways in which a secret may be produced and preserved. In the final section 8 I propose to classify such ‘secretization strategies’ in four categories:

- (a) Hampering of access to the information (by hiding or coding it).
- (b) Proliferation of false information (about the secret or around it).
- (c) Production of uncertainty (indefeasible defeaters).
- (d) Overload of information (excess of unessential data).

Even if these strategies are expressed in terms of information, they support the main claim of this paper, namely that secret is the frustration of knowledge, not the concealment of information.

Works cited

- [1] Ahmed, Sara. (2010). “Secrets and silence in feminist research”. In *Secrecy and silence in the research process: feminist reflections*, Róisín Ryan-Flood and Rosalind Gill, eds. Abingdon, Oxon ; New York, NY: Routledge.
- [2] Barton, Alana, y Howard Davis. (2018). *Ignorance, Power and Harm Agnotology and The Criminological Imagination*. Cham: Springer International Publishing.
- [3] Birchall, Clare, ed. (2011). “Monographic section on ‘Secrecy and Transparency’”. *Theory, Culture and Society* 28 (7-8): 7-25.
- [4] — (2014). “Aesthetics of the Secret”. *New Formations* 83 (83): 25-46.
- [5] — (2016). “Managing Secrecy”, 13.
- [6] Borradori, Giovanna. (2016). “Between Transparency and Surveillance: Politics of the Secret”. *Philosophy & Social Criticism* 42 (4-5): 456-64.
- [7] Broncano, Fernando. (2019). *Puntos ciegos: ignorancia pública y conocimiento privado*. Madrid: Lengua de Trapo.
- [8] — (2020). *Conocimiento expropiado: epistemología política en una democracia radical*. Madrid: Akal.
- [9] Gross, Matthias, y Linsey McGoey, eds. (2015). *Routledge International Handbook of Ignorance Studies*. London ; New York: Routledge, Taylor & Francis Group.

- [10] Mills, Charles W. (1997). *The Racial Contract*. Ithaca: Cornell University Press.
- [11] Peels, Rik, y Martijn Blaauw, eds. (2016). *The epistemic dimensions of ignorance*. New York: Cambridge University Press.
- [12] Proctor, Robert, y Londa L. Schiebinger, eds. (2008). *Agnotology: The Making and Unmaking of Ignorance*. Stanford, Calif: Stanford University Press.
- [13] Resnik, David B. (2006). "Openness versus Secrecy in Scientific Research". *Episteme* 2 (3): 135-47.
- [14] Sullivan, Shannon, y Nancy Tuana, eds. (2007). *Race and Epistemologies of Ignorance*. Albany: State University of New York Press.

Yet another defence of doxasticism about delusions

Miguel Núñez de Prado Gordillo*
Universidad Autónoma de Madrid

Manuel Almagro Holgado**
Universidad de Granada

Delusions are typically conceived in terms of beliefs in the scientific and clinical literature. A particularly relevant exemplar of this can be found in the DSM-5, where delusions are defined as “fixed *beliefs* that are not amenable to change in light of conflicting evidence” (APA 2013: 87, emphasis added). However, this standard doxasticist approach has not passed uncriticized and there is an ongoing debate regarding its viability. On the one hand, several authors have criticized it on the grounds that delusional states do not properly meet the necessary criteria that certain theories of belief (namely, interpretivism and functionalism) establish for true or appropriate belief ascription (Currie 2000, Egan 2008, Porcher 2019, Schwitzgebel 2012, Tumulty 2012). Specifically, these anti-doxasticist approaches deny belief-status to delusions on the grounds that many people with delusions exhibit attitude-attitude and/or attitude- behavioral inconsistencies (i.e., deviations from stereotypical belief-like causal roles or rationality constraints).

By contrast, some authors have vindicated the doxastic approach to delusional states, either on the grounds that a) functionalism and interpretivism still grant a doxastic status to delusions when certain considerations are taken into account; or b) functionalism and interpretivism are inadequate theories of belief and should thus be replaced by a different theoretical framework. As examples of the former *revisionist* strategies, Bortolotti (2010, 2012) has advocated for a modest interpretationist defence of doxasticism, while Bayne & Pacherie (2005) have defended a functionalist variety of doxasticism based on Schwitzgebel’s (2013) dispositionalist account. As a recent example of the latter strategy, Clutton (2018) has rejected both interpretivism and functionalism because of their alleged ‘anti-realist tendencies’ towards belief. He has proposed instead a cognitive- phenomenological defense of scientific doxasticism, based on the idea that beliefs are real entities, neural states whose triggering produce in the subject a distinctive kind of cognitive phenomenological experience: that of mentally assenting to p when p - entertaining triggers obtain.

* miguel.nunnezdeprado@uam.es

** malmagro@ugr.es

These pro-doxasticist approaches typically draw from two different desiderata: a) a *scientific desideratum*, related to the claim that doxasticism leaves us in a better position to account for the causal processes involved in the acquisition, adoption and persistence of delusional states (which might thus in turn have implications for the comprehension or development of our psychiatric and psychological intervention designs and techniques); and b) an *ethical-political desideratum*, related to the claim that doxasticism is best suited for defending the idea that delusional phenomena can be understood in intentional or intelligible terms, thus safeguarding our ascriptions of agency, responsibility and autonomy to the person diagnosed. This, in turn, would provide a further barrier protection for their right to ethical treatment (including things like informed consent, employment of non-invasive intervention techniques, etc.).

Our main goal in this paper will be to show why the above-mentioned pro-doxastic proposals cannot meet both desiderata at the same time; specifically, revisionist proposals cannot meet the scientific desideratum, while Clutton's cognitive-phenomenological approach cannot meet the ethical-political desideratum. This, we argue, is due to certain problems regarding the notion of belief at play in each pro-doxasticist approach.

On the one hand, we will argue that the reason why revisionist proposals fall short of their scientific aspirations is that, in order to secure a doxastic understanding of delusional phenomena, they need to implement an ascriber-relativist approach to the truth of belief ascriptions (Tumulty 2011). However, if the truth value of belief ascriptions is relative to the ascriber's standards and communicative interests, then the revisionist's notion of belief can hardly constitute an adequate explanatory tool for a scientific account of delusions and their clinical intervention. Some authors have drawn from a similar argument to vindicate anti-doxasticism towards delusions (Schwitzgebel 2012, Tumulty 2011). By contrast, we will argue that doxasticism can and should be primarily defended on the grounds of its ethical and political benefits.

On the other hand, Clutton's proposal can be construed as an attempt to avoid the ascriber-relativism of revisionist proposals. He establishes that the "truth-marker" of a belief ascription is the neural state that realizes the belief in question, i.e., the cognitive-phenomenological disposition to mentally assent to p when p -entertaining triggers obtain. However, we will argue that this operationalized cognitive scientific notion of belief lacks the normative force that characterizes folk-psychological belief ascriptions. Therefore, Clutton's cognitive phenomenological defense ends up falling short of the ethical-political aspirations of doxasticism.

We will call this the objection from Bifurcation since, as matters stand, we seem to come to a crossroads: we can either take sides with revisionist doxasticism, which preserves the folk psychological notion of belief, or take sides with Clutton's scientific doxasticism, committing ourselves to the project of redefining the notion of belief in terms appropriate for scientific description and explanation purposes. While the former would

imply to forego the scientific desideratum, the later would imply to forego the ethical-political desideratum.

We will argue that the problem of Bifurcation only follows from an incompatibilist approach to belief ascription, according to which there can only be one notion of belief in place (presumably, one that is the most informative for scientific research), to which delusions must conform in order to vindicate their doxastic status. Instead, we will advocate for a compatibilist approach, which draws from the recognition that there are two notions of belief in place in the debate on the doxastic status of delusions.

Sánchez Curry (2020) has recently proposed a similar approach to belief ascription. He has distinguished between what he calls *attitudes of belief* (i.e., the folk psychological notion of belief) and *cognitive states of belief* (the notion of belief deployed in cognitive scientific research). From this compatibilist point of view, we are no longer forced to choose between preserving the folk psychological notion of belief or redefining it in purely descriptive terms that can be of use for scientific investigation. We can just accept the bifurcation, or, in other words that there are two uses of the concept of ‘belief’ that are discontinuous with each other. Each are individuated by the different social practices in which they are employed: a) on the one hand, ascriptions of attitudes of belief typically take part in normative, value-laden accounts of the affairs of different living beings, aimed at the goal of rationalizing, justifying and assessing their doings with regard to certain intersubjective social standards (of rationality, morality, psychological well-being, etc.); b) on the other hand, ascriptions of cognitive states of belief play their role in nomological accounts of the affairs of different living beings, aimed at their precise causal explanation, prediction and control. Consequently, the truth conditions of ascriptions of attitudes of belief are independent from the truth conditions of ascriptions of cognitive states; whether one can be truthfully ascribed the attitude of belief that *p* is independent from whether one can be truthfully ascribed the cognitive state of belief that *p*, and *vice versa*.

Once this is established, we can clearly distinguish between two kinds of doxasticism towards delusions: a) folk-psychological doxasticism, according to which delusions are properly understood in terms of attitudes of belief; and b) cognitive scientific doxasticism, according to which delusions are properly understood in terms of cognitive states of belief. We argue in favor of folk psychological doxasticism about delusions on the grounds of its ethical and political value, while remaining agnostic regarding cognitive scientific doxasticism. Finally, we conclude by suggesting that revisionist doxasticism can partner up well with a non-descriptivist view of folk psychological belief ascription practices (Pérez-Navarro et al. 2019, Sánchez-Curry 2020). This approach can provide the kind of ascriber-relativist approach to the truth of belief ascriptions that revisionist doxasticism needs to grant a doxastic status to delusions. In addition, it allows us to bring to the forefront the ethical and political motivations behind doxasticism.

References

- [1] American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub.
- [2] Bayne, T., & Pacherie, E. (2005). In defence of the doxastic conception of delusions. *Mind & Language*, 20(2), 163-188.
- [3] Bortolotti, L. (2010). *Delusions and other irrational beliefs*. Oxford University Press.
Bortolotti, L. (2012). In defence of modest doxasticism about delusions. *Neuroethics*, 5(1), 39-53.
- [4] Clutton, P. (2018). A new defence of doxasticism about delusions: The cognitive phenomenological defence. *Mind & Language*, 33(2), 198-217.
- [5] Currie, G. (2000). Imagination, delusion and hallucinations. *Mind & language*, 15(1), 168-183.
- [6] Egan, A. (2008). Imagination, delusion, and self-deception. In T. Bayne & J. Fernandez (eds.), *Delusion and Self-Deception: Affective and Motivational Influences on Belief Formation (Macquarie Monographs in Cognitive Science)* (pp. 263-280). Psychology Press.
- [7] Pérez-Navarro, E., Fernández Castro, V., González de Prado Salas, J., & Heras–Escribano, M. (2019). Not expressivist enough: Normative disagreement about belief attribution. *Res Philosophica*.
- [8] Porcher, J. E. (2019). Double bookkeeping and doxasticism about delusion. *Philosophy, Psychiatry, & Psychology*, 26(2), 111-119.
- [9] Sánchez-Curry, D. (2020). Interpretivism and norms. *Philosophical Studies*, 177(4), 905-930.
- [10] Schwitzgebel, E. (2012). Mad belief? *Neuroethics*, 5(1), 13-17.
- [11] Schwitzgebel, E. (2013). A dispositional approach to attitudes: Thinking outside of the belief box. In N. Nottelman (ed.) *New essays on belief* (pp. 75-99). Palgrave Macmillan, London.
- [12] Tumulty, M. (2011). Delusions and dispositionalism about belief. *Mind & language*, 26(5), 596-628.
- [13] Tumulty, M. (2012). Delusions and not-quite-beliefs. *Neuroethics*, 5(1), 29-37.

Etiología doxástica y proposiciones lotería¹

Manuel Pérez Otero*

Departamento de Filosofía. Facultad de Filosofía.

Universidad de Barcelona

LOGOS (Research Group in Analytic Philosophy)**

Una *proposición lotería* establece –de forma explícita o implícita– que no saldrá, o no salió, cierto número particular en una lotería o un sorteo de características similares. Por ejemplo, cuando el contexto permite remitirnos con claridad a una lotería concreta con un millón de posibles resultados diferentes, el enunciado ‘no saldrá el 597’ expresa una proposición lotería. Restrinjo el estudio de estas proposiciones a los casos que albergan mayor interés epistémico: cuando la base para creer dicha proposición es meramente el razonamiento probabilístico, no –por ejemplo– la consulta de un periódico o tener indicios de que el sorteo esté sesgado.

Es muy generalizada la intuición de que no tenemos conocimiento de una proposición lotería verdadera (en las condiciones indicadas; cuando sólo el cálculo probabilístico nos permitiría creerla), incluso si –esto sería controvertido– tenemos justificación para creerla. Comparto la intuición y también el juicio que dicha intuición respalda: no hay saber en ese tipo de situaciones. Quien rechace que conozcamos proposiciones lotería puede tener problemas si cree que existe saber, o al menos suspende el juicio, en ciertos otros casos muy similares a las loterías pero no consigue explicar la diferencia relevante.

Considero poco apropiado el uso laxo de Hawthorne (2004) de la etiqueta “proposiciones estilo Vogel sobre loterías”, precisamente porque incluye como tales algunos ejemplos bastante diferentes a las loterías, para los cuales no es tan apremiante el reto de explicar la diferencia relevante mencionada; uno de esos ejemplos, sobre la posibilidad de que nos hayan robado el coche, procede de Vogel (1990). Cosa bien distinta es el caso *Bad Apple*, ideado por E. Sosa (cf. Sosa 2015, cap. 5). Supongamos que la compañía Apple tiene un millón de operadores telefónicos que responden preguntas sobre Apple de cualquier usuario. Uno de ellos proporciona siempre una respuesta falsa, pero todos los otros responden correctamente con absoluta fiabilidad. Cuando alguien llama para hacer su consulta, el sistema selecciona aleatoriamente a uno de los operadores; y todos tienen

* perez.otero@ub.edu

** http://www.ub.edu/grc_logos/manuel-perez-otero

¹ Financiación: Proyecto “*Etiología doxástica, conceptualizaciones singulares y la naturaleza de la filosofía*” (PID2019-105728GB-I00), (MINECO/FEDER, UE) / Grupo de investigación consolidado *LOGOS* (2017SGR63), AGAUR.

la misma probabilidad de ser seleccionados. S conoce esos datos. Si S obtiene una respuesta positiva y verdadera tras llamar preguntando por la proposición Q, ¿sabe S que Q es verdad? Según Sosa, nuestra intuición favorece atribuir conocimiento a S. Y, de forma muy excepcional, propone que también hay saber respecto a proposiciones lotería verdaderas.² Muchos tenemos reticencias a negar el saber en el caso *Bad Apple*, incluso si no compartimos la intuición de Sosa y nos limitamos a suspender el juicio. Si preservamos la intuición previa sobre las proposiciones lotería, constatamos el reto teórico de explicar dónde podría residir la diferencia relevante. Ése es el *problema Bad Apple*.

Epistemólogos contemporáneos debaten sobre si el saber y/o la justificación epistémica requieren que la creencia correspondiente sea *segura* [*safe*], en el sentido correspondiente a SEG:

SEG: *si el sujeto hubiera tenido la creencia, ésta habría sido verdadera*

Como sucede con otros conceptos técnicos (fiabilidad, sensibilidad) esa noción suele definirse en virtud de lo que sucedería en *otros* mundos posibles cercanos al mundo real; pero dicha definición (en esa versión estándar) involucra la creencia original. Cuando ésta es una creencia de carácter *singular* (cuya individuación depende de una entidad concreta), la definición resulta entonces poco útil, pues en algunos de esos otros mundos posibles la creencia relevante resulta ser *otra* (puede suceder eso, por ejemplo, si la creencia original es sobre una oveja percibida ubicada en un rebaño de muchas otras ovejas semejantes). Conceptos como fiabilidad, sensibilidad y seguridad se redefinen a veces en función de la etiología doxástica; es decir, aunque nos centremos en una creencia singular, C, podemos manejar una definición apropiada que atienda a los resultados del proceso que ha conducido a C en el mundo real, incluso respecto a mundos posibles donde el resultado es una creencia (relevantemente semejante aunque) diferente a C. Tanto Pritchard (2005, p. 163) como Sosa (1999) mencionan *métodos* en sus teorías; pero no consiguen evitar la dificultad a la que nos referimos, derivada del carácter singular de ciertas creencias. Lasonen Aarnio (2010), por el contrario, emplea una definición de creencia segura en la cual interviene de forma relevante la etiología del proceso doxástico: la creencia C es segura si y sólo si el método que la produjo no podría fácilmente haber producido una creencia falsa. Sin necesidad de asumir que una proposición lotería sea singular, hemos explorado si una noción de seguridad en la cual resulta así relevante la etiología doxástica permite solucionar el problema *Bad Apple*.

El resultado de la investigación es negativo. La diversidad de procesos cuyo resultado es la creencia y la diversidad de variables que podrían ser relevantes en los casos lotería y *Bad Apple*, conllevan varios modos posibles de determinar si en esos casos hay

² Sosa sostiene que hay saber en los casos lotería no sólo por su semejanza con el caso *Bad Apple*, sino porque cree que sus definiciones del conocimiento así lo implican. A mi juicio, sin embargo, esas definiciones son neutrales sobre casos lotería; cf. Pérez Otero (2016).

seguridad. Pero ninguno permite decir con rotundidad que haya seguridad en el caso Bad Apple pero inseguridad en el caso lotería (excepto si existe otra variable relevante que no hemos sabido identificar).

A continuación, presento muy someramente algunos detalles relevantes de dicha investigación.

El proceso doxástico P (que ha tenido como resultado cierta creencia) es seguro *si P hubiera ocurrido, habría producido una creencia verdadera*. Cuando su antecedente, Q, es verdadero, resulta difícil interpretar un enunciado del tipo *Si Q hubiera sido el caso, R habría sido el caso*. No es claro qué elemento debe variar al considerar mundos posibles cercanos diferentes al mundo real. Por ello, las definiciones de creencia segura y proceso doxástico seguro se aplican con dificultades. Adicionalmente, tenemos el problema de cómo individualizar el proceso doxástico (semejante al *problema de la generalidad* que afecta al fiabilismo).

El caso lotería involucra dos variables relevantes: la compra-selección del boleto; el sorteo. Bad Apple involucra tres variables relevantes: la asignación-selección aleatoria del operador telefónico; el valor de verdad de la proposición Q resultante de la consulta; qué operador telefónico es el que miente. La creencia-lotería no es segura (conforme a la definición SEG). En algunos mundos posibles cercanos sale el 597. Por tanto, “*Si S hubiera creído ‘no saldrá el 597’, la creencia habría sido verdadera*” es falso. No es claro si la creencia-Apple es segura: Si varía (respecto al mundo real) el número que Apple asigna, entonces es segura (la variación no será que asignen el número mentiroso, pues entonces S no tendría la creencia). Si varía el operador telefónico mentiroso, entonces es segura (el número asignado no será mentiroso, pues entonces S no tendría la creencia). Si varía el valor de verdad de Q, entonces no es segura (S, de todos modos, creería Q). Esa indeterminación se correlacionaría con la indeterminación de nuestras intuiciones sobre si la creencia-Apple constituye conocimiento.

En ninguno de los dos casos (lotería; Bad Apple), el proceso doxástico es seguro. ¿Hasta dónde llega el proceso doxástico, en casos lotería y en casos Bad Apple? Respecto a la lotería: Si el proceso incluye la compra-selección del boleto, entonces no es seguro (en algunos mundos posibles cercanos sale el 597). Si el proceso no incluye la compra-selección del boleto, entonces no es seguro (incluso si fijamos el número premiado, quizá el boleto seleccionado –diferente al 597– sea premiado). Respecto a Bad Apple: Si el proceso incluye-fija el número que Apple asigna y que ese número es sincero, y lo que varía es el valor de verdad de Q, entonces no es seguro (el sujeto creará Q). Si el proceso incluye-fija qué operador miente y el valor de verdad de Q, y varía el número que Apple asigna, entonces no es seguro (quizá Apple asigne el número mentiroso). Si el proceso incluye-fija el número que Apple asigna y el valor de verdad de Q, y varía qué operador miente, entonces no es seguro (quizá ese número asignado miente). En Bad Apple, *el proceso podría ser seguro si existiera otro factor relevante que pudiera variar* (quedando fijadas las otras tres variables) teniendo como resultado que la creencia fuera verdadera.

Bibliografía

- [1] Hawthorne, John (2004): *Knowledge and Lotteries*, Oxford: Oxford University Press.
- [2] Lasonen Aarnio, Maria (2010): “Unreasonable Knowledge”, *Philosophical Perspectives* 24, pp. 1-21.
- [3] Pérez Otero, Manuel (2016): “La teoría competencial del saber de E. Sosa”, *Teorema* 35/2, pp. 181-195. Reimpreso en M. Pérez Otero, *Ensayos filosóficos sobre el conocimiento y el escepticismo epistémico*, Saarbrücken: Editorial Académica Española, 2017.
- [4] Pritchard, Duncan (2005): *Epistemic Luck*, Oxford: Oxford University Press.
- [5] Sosa, Ernest (1999): “How to defeat opposition to Moore”, en J. Tomberlin (ed.), *Philosophical Perspectives 13: Epistemology*, Blackwell, pp. 141–54.
- [6] Sosa, Ernest (2015): *Judgment and Agency*, Oxford: Oxford University Press.
- [7] Vogel, Jonathan (1990): “Are there Counterexamples to the Closure Principle?”, en M. Roth y G. Ross (eds.), *Doubting: Contemporary Perspectives on Skepticism*, Dordrecht: Kluwer.

Collective Understanding-Why

Dani Pino*
University of Seville

In this paper I provide an account for collective understanding-why in three steps. First, I show what understanding-why consists of by framing it within the problem of reductionism of understanding-why. This problem involves a debate about the kind of abilities that are required in order to understand why $\neg p$, for which I will put forward a solution appealing to upstream and downstream abilities, i.e., abilities in order to acquire and manifest understanding, respectively. Second, I consider the case of understanding-why in the context of Collective or Group Epistemology, which is the branch of Epistemology that considers groups as subjects whose knowledge is not reducible to its individual members' knowledge, nor results from the summation of them. In order to clarify the way groups may understand why p , I will make an analogy with a well-known case of collective cognitive agency, that is, transactive memory—a phenomenon according to which a collection of two or more individuals may encode, store and retrieve memories in such a way that any of them are able to do in isolation. Finally, I show an example of collective understanding-why that is quite familiar to epistemologists: the central case of hermeneutical injustice (Fricker 2007, 149-150).

In the first of these three steps I distinguish understanding-why from other instances of understanding. The debates on the nature of understanding identify three kinds of phenomena. Firstly, when we say “I understand that it is important to have a balanced diet to get healthy”, the object of understanding is a proposition, so we talk of propositional understanding—a subject S understands that p . Secondly, we can understand a whole subject matter, so we could say something like “I understand Spinoza's notion of a composed individual”. This is a case of objectual understanding, where S understands p . Finally, someone might say “I understand why my TV is not working”, which is an instance of understanding-why— S understands why p . In this paper, I will focus on the third sort of understanding.

There is not unanimous agreement about the relationship between these three kinds of understanding. Be that as it may, I will not delve into those debates. Instead, I will take understanding-why in isolation as a particular instance of understanding regardless whether it is in fact tantamount to propositional understanding or objectual understanding. This is not an uncontroversial kind though, for there is no agreement about the nature of understanding-why either. The debate is focused on whether understanding why may be

* dps@us.es

reducible to knowledge-why or not. Reductionists claim that it is indeed, so since knowledge-why is also reducible to knowledge that, the object of understanding is a proposition or a set of propositions that maintain a causal relationship with each other. Thus, reductionist approach may be express as follows:

REDUCTIONISM:

S understands why p iff S knows that p is because of q.

Reductionism hence claims that understanding-why is constituted by the subject's knowledge of two propositions that hold a causal relationship between each other. Against this approach, non-reductionists deny that understanding why may be reducible to knowledge why. For them, the object of this phenomenon is a grasp whose object is the relationship itself that makes sense that p because q. If I say that I understand why Sara was late (p), and she was late because she was in a traffic jam (q), the reductionist focuses on the explanation of why p, so that q suffices in order to understand p. Meanwhile, the non-reductionist pays attention to how p and q relate, not just to the mere propositions. Thus, for the non-reductionist what really matters is her grasping that relationship, i.e., understanding-why amounts to grasp the relationship between p and q. We can then flesh it out like this:

NON-REDUCTIONISM:

S understands why p iff S grasps the relationship between q as the cause of p.

The distinction between both accounts is accurately mirrored in the Hills-Sullivan debate on the nature of understanding-why. The problem is firstly set by Allison Hills (2015), who claims that grasping amounts to having the relationship between propositions under the understander's control. But what does it consist of having that relationship under one's control? Hills says that when understanding why p, one can exercise a set of abilities required in order to manipulate the relationship in which the grasp consists of. On her view, it amounts to have the relevant know-how to control the way p and its explanatory reason q relate. This set of six abilities constitute what she coins as cognitive control.

In the opposite corner of the ring, Emily Sullivan (2018), denies that the abilities involved in understanding-why are akin to know-how abilities. According to her, Hills misunderstands the nature of the abilities that leads to the formation of a grasp. Sullivan does not discuss the nature of the grasp itself—she rather focuses on underlining that know-how requires a sort of ability, specifically embodied ability, that differs from those required in order to form a grasp in the terms that Sullivan conceives the abilities that constitutes the cognitive control.

I hold that Hills and Sullivan are talking past each other, for one is taking the abilities involved in understanding-why as upstream abilities—that is, the kind of abilities that

are exerted in order to acquire understanding—and the other is assuming that understanding-why requires the performance of downstream abilities—that is, abilities that are exerted in order to manifest understanding. Accordingly, they have not contending views, but rather complementary views.

Once established, for the second step of my argument I consider the case of understanding-why in collective agents, that is, collections of individuals that act as a whole system whose actions are not reducible to, nor result from, the mere summation of the members' activity. Rather, the activity of the whole system results from the distribution of actions—either cognitive (Bird 2010, Palermos 2016) or practical (Palermos and Tollefsen 2018, Birch 2018)—among the individual members. In order to build in the case, I will bring forward transactive memory as a model to understand the nature of collective understanding.

Transactive memory is accepted to be a sound evidence in favour of the existence of collective subjects and group agency (Tollefsen et al. 2013). According to the description of the phenomenon, a system formed by two or more individuals may be able to encode, store and retrieve information in such a way that it would not be accessible for each individual in isolation (Huebner 2016, Wegner 1987). The explanation for this is that some feedback dynamics occur due to the interaction between individuals, which is responsible for the cognitive activity that amounts to memory. Therefore, what is in play is the activity of the system as a whole, not merely as an aggregation of individuals. My hypothesis is that there is a distribution of upstream and downstream abilities among the individual members of the system that contributes to the overall state of grasping that amounts to a strictly collective understanding-why.

The third and final step of the paper exemplifies my account for collective understanding-why in a well-known case within Epistemology: the central case of hermeneutical injustice (Fricker 2007, 149-150). According to this case, a hermeneutical lacuna is shown by a collection of women that are unable to understand the disgusting experience they individually feel due to having been victims of what they, after working cognitively together, label as sexual harassment. If my hypothesis succeeds, Fricker's case could be explained in terms of a distribution of the abilities that are required in order to understand why the women that were victims of sexual harassment had that common experience they could not individually explain. This sort of distribution parallels the way some cases of collective knowledge are produced by the distribution of cognitive or practical resources among the individual members of a group epistemic agent. Therefore, we could be in dealing with a novelty phenomenon within Group Epistemology that it would be worth further inquiring.

References

- [1] Birch, J. (2018). "Joint Know-How". *Philosophical Studies*, 1-24.
- [2] Bird, A. (2010). "Social Knowing: The Social Sense of 'Scientific Knowledge'". *Philosophical Perspectives*, 24(1), 23-56.
- [3] Hills, A. (2015). "Understanding Why". *Noûs*, 49(2):661-688.
- [4] Huebner, B. (2016). "Transactive Memory Reconstructed: Rethinking Wegner's Research Program". *Southern Journal of Philosophy*, 54(1):48-69.
- [5] Fricker, M. (2007). *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford University Press.
- [6] Palermos, S. O. (2016). "The Dynamics of Group Cognition". *Minds and Machines*, 26(4), 409-440. Palermos, S. O., & Tollefsen, D. P. (2018). "Group Know-How". In J. A. Carter, A. Clark, J. Kallestrup, S. O. Palermos, & D. Pritchard (Eds.), *Socially Extended Epistemology* (pp. 112-131). Oxford University Press.
- [7] Sullivan, E. (2018). "Understanding: Not Know-How". *Philosophical Studies*, 175(1):221-240.
- [8] Tollefsen, D. P., Dale, R., & Paxton, A. (2013). "Alignment, Transactive Memory, and Collective Cognitive Systems". *Review of Philosophy and Psychology*, 4(1):49-64.
- [9] Wegner, D. M. (1987). "Transactive Memory: A Contemporary Analysis of the Group Mind". In B. Mullen & G. R. Goethals (Eds.), *Theories of Group Behavior* (pp. 185-208). Springer.

Painful Moods and the Affective Component of Primary Dysmenorrhea

Carlota Serrahima*
Universitat de Barcelona

Manolo Martínez**
Universitat de Barcelona

The usual target of the study of pain in analytic philosophy of mind is the discrete episode of nociception, typically caused by tissue damage. Usual examples involve acute thermal pain caused by scalding water (Bain, 2013), dull aches caused by fractured ankles (Klein, 2015) or pain in one's forearm caused by "the presence or damage or disturbance of tissue in a forearm" (Cutter & Tye, 2011, p. 92). While this kind of pain is undoubtedly important and theoretically interesting, focus on it has had the effect of pushing other kinds of pains (some of which, arguably, with more sustained and substantial impact on the health and wellbeing of their sufferers) towards the periphery of the philosophical debate on pain. It has also resulted in an implicit taxonomy of affective mental states which sharply distinguishes pains from moods.

In this paper, we focus on the affective component of primary dysmenorrhea (that is, menstrual pain in the absence of additional clinical conditions). This is a central kind of pain -- in that it is epidemiologically extremely common (between 45 and 95% of menstruating women suffer it, Iacovides et al., 2015), and in that it demonstrably impacts the wellbeing of its sufferers in a very substantial way (Bajalan et al., 2018; Iacovides et al., 2014). Philosophical discussion of this condition (which has so far been all but completely ignored) is, we believe, long overdue, not only because of its sheer prevalence, but also because it problematizes the taxonomy of affective states that, we suggest, philosophers implicitly rely on. We argue, on biological and phenomenological grounds, that dysmenorrhea cannot be simply analyzed as a mere composite of pain- and mood-related ingredients -- even if it does include them.

On the one hand, pain is undoubtedly a very substantial component of the suffering related to dysmenorrhea. This pain component is philosophically interesting in and of itself. First, because of its temporal profile, which is neither acute nor chronic: it occurs cyclically (Przybylo & Fahs, 2018) in relatively predictable episodes. This makes menstrual pain sometimes lead to central sensitization (Granot et al., 2001) with a host of debilitating

* carlotaserrahima@gmail.com

** manolo@austrohungaro.com

consequences, as happens with chronic pains. Yet, unlike chronic pains, those episodes are relatively short-lived, and are often assumed to be part of normal bodily function (Wilson & Keye, 1989, p. 320). Hence we speculate, drawing also on personal experience, that this makes them less prone to the kind of catastrophizing that usually accompanies long-term pains (pace Tousignant-Laflamme & Marchand, 2009). Second, dysmenorrheic pain is interesting because the connection of menstrual cramps to maladaptiveness (which lies at the core of many prominent philosophical accounts of painfulness) is at the very least complicated: the kind of muscular contractions that give rise to the cramps are necessary in order to dispose of endometrial tissue (Smith, 2018 p. 3). This means that so-called “tracking”, or teleosemantic brands of representationalism (Bain, 2013; Cutter & Tye, 2011; Tye, 2006) have to provide slightly more complicated accounts of how this mostly adaptive kind of contractions lead to, e.g., representations of bodily events as bad -- which is what many of them propose as the representationalist analysis of painfulness. Finally, given the scarcity of location-related information coming from viscerae, menstrual cramps are usually referred to other parts of the body, such the lower back or thighs. This results in an imprecisely located (and also, for representationalists, illusorily located) pain experience. None of these characteristics are impossible to accommodate in leading contemporary accounts of painfulness. However, all of them are exceptions and problem cases in most such accounts.

On the other hand, and importantly, dysmenorrhea does not just involve pain, but also a cluster of mood-related phenomena. There are very robust results linking dysmenorrhea to anxiety, depression and stress (Bajalan et al., 2018 is a recent review). Women with severe primary dysmenorrhea also have significantly lower scores in the quality of life enjoyment and satisfaction questionnaire (Iacovides et al., 2014). Besides, and most importantly for our purposes, standard questionnaires of distress related to dysmenorrhea (Moos, 1968) include items concerning what philosophers would call mood (e.g., loneliness, anxiety, restlessness, irritability, depression, etc; op. cit. p. 857), apart from other items having to do with cognitive symptoms (e.g., forgetfulness, confusion, lowered judgement, etc.; *ibid.*). In other words, part of the standard clinical conception of dysmenorrhea has to do, not so much with physical disturbances, but with emotional ones.

Philosophical discussion on affective states sharply distinguishes between pains and moods (e.g, Bordini, 2017; Kind, 2013; Rossi, 2019). However, in our view, dysmenorrhea doesn't sit comfortably in this taxonomy: pain and mood components of dysmenorrhea are intertwined in a particularly intimate way. First, we contend, dysmenorrhea is not just the juxtaposition of a pain and a mood. Enduring dysmenorrhea is not at all like feeling a cut on one's left pinky while being sad that the spring break is over. The two components appear to speak to each other, sufferers describing their experience, for instance, thus: “I would say, when the period's just about to start, or just before, you know, within an hour before, I feel tearful ... It's as if my tears alert me” (testimonial from Burbeck & Willig, 2014). Second, it is also probably not just that the mood disorder is caused by the

pain. There is neurophysiological evidence that both of them are more intimately related. For example, the top-down component of pain (which is in turn connected to central sensitization, as described above) and depressive moods share biological pathways and neurotransmitters (Bair et al., 2003). In other words, it is not clearly the case that having one causes the other, but rather that part of what it is to have one is to have the other. Some researchers talk of a pain-depression dyad to capture this intimate relation (Goldenberg, 2010). We suggest that dysmenorrhea is a prime example of this unified painful-mood state.

In conclusion, dysmenorrhea appears as an atypical phenomenon in several prominent, mostly independent debates in the philosophy of affective phenomenology. If only because of its extremely high prevalence, this is an uncomfortable situation for the field to be in. Our piece is an invitation for philosophers to give dysmenorrhea its due place in the investigation of affective phenomenology.

References

- [1] Bain, D. (2013). What Makes Pains Unpleasant? *Philosophical Studies*, 166, S69-S89.
- [2] Bair, M., Robinson, R., & Kroenke, K. (2003). Bair MJ, Robinson RL, Katon W, Kroenke K. *Depression and pain comorbidity: A literature review. Arch Intern Med* 163: 2433-2445. *Archives of Internal Medicine*, 163, 2433–2445. <https://doi.org/10.1001/archinte.163.20.2433>
- [3] Bajalan, Z., Moafi, F., MoradiBaglooei, M., & Alimoradi, Z. (2018). Mental health and primary dysmenorrhea: A systematic review. *Journal of Psychosomatic Obstetrics & Gynecology*. <https://www-tandfonline-com.sire.ub.edu/doi/abs/10.1080/0167482X.2018.1470619>
- [4] Bordini, D. (2017). Not in the Mood for Intentionalism. *Midwest Studies In Philosophy*, 41(1), 60–81. <https://doi.org/10.1111/misp.12066>
- [5] Burbeck, R., & Willig, C. (2014). The personal experience of dysmenorrhoea: An interpretative phenomenological analysis. *Journal of Health Psychology*, 19(10), 1334–1344. <https://doi.org/10.1177/1359105313490313>
- [6] Cutter, B., & Tye, M. (2011). Tracking Representationalism and the Painfulness of Pain. *Philosophical Issues*, 21, The Epistemology of Perception.
- [7] Goldenberg, D. L. (2010). Pain/Depression Dyad: A Key to a Better Understanding and Treatment of Functional Somatic Syndromes. *The American Journal of Medicine*, 123(8), 675–682. <https://doi.org/10.1016/j.amjmed.2010.01.014>

- [8] Granot, M., Yarnitsky, D., Itskovitz-Eldor, J., Granovsky, Y., Peer, E., & Zimmer, E. Z. (2001). *Pain Perception in Women With Dysmenorrhea*. 98(3), 5.
- [9] Iacovides, S., Avidon, I., & Baker, F. C. (2015). What we know about primary dysmenorrhea today: *A critical review*. *Human Reproduction Update*, 21(6), 762–778.
- [10] Iacovides, S., Avidon, I., Bentley, A., & Baker, F. C. (2014). Reduced quality of life when experiencing menstrual pain in women with primary dysmenorrhea. *Acta Obstetrica et Gynecologica Scandinavica*, 93(2), 213–217.
- [11] Kind, A. (2013). The Case Against Representationalism About Moods. In U. Kriegel (Ed.), *Current Controversies in Philosophy of Mind*.
- [12] Klein, C. (2015). What the body commands: *The imperative theory of pain*. MIT Press.
- [13] Moos, R. H. (1968). The Development of a Menstrual Distress Questionnaire. *Psychosomatic medicine*, 15.
- [14] Przybylo, E., & Fahs, B. (2018). Feels and Flows: On the Realness of Menstrual Pain and Crippling Menstrual Chronicity. *Feminist Formations*, 30(1), 206-229. <https://doi.org/10.1353/ff.2018.0010>
- [15] Rossi, M. (2019). A perceptual theory of moods. *Synthese*. <https://doi.org/10.1007/s11229-019-02513-1>
- [16] Smith, R. P. (2018). *Dysmenorrhea and Menorrhagia*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-71964-1>
- [17] Tousignant-Laflamme, Y., & Marchand, S. (2009). Excitatory and inhibitory pain mechanisms during the menstrual cycle in healthy women. *Pain*, 146, 47–55. <https://doi.org/10.1016/j.pain.2009.06.018>
- [18] Tye, M. (2006). Another Look At Representationalism About Pain. In M. Aydede (Ed.), *Pain: Ew Essays on Its Nature and the Methodology of ts Study* (pp. 99–120). The MIT Press. Bradford Books.
- [19] Wilson, C. A., & Keye, W. R. (1989). A survey of adolescent dysmenorrhea and premenstrual symptom frequency: A model program for prevention, detection and treatment. *Journal of Adolescent Health Care*, 10(4), 317–322. [https://doi.org/10.1016/0197-0070\(89\)90065-X](https://doi.org/10.1016/0197-0070(89)90065-X)

Epistemología del desinterés por la verdad

Pablo Vera Vega*
Universidad de La Laguna

El objetivo de esta ponencia es la crítica del concepto de bullshit en tanto que derivado de una cierta actitud epistémicamente viciosa, que es la insouciance. El fenómeno de la bullshit, a pesar de parecer más sociológico que filosófico, ha sido ampliamente estudiado desde el ámbito de la Filosofía. Por bullshit se entiende desde la tontería hasta la sandez pasando por prácticamente todo el campo semántico de la mentira. Sin embargo, en los últimos años, este concepto se ha tecnificado dentro de la reflexión filosófica. Fue Harry Frankfurt, con su *On Bullshit*, en su reedición del año 2005, quien logró que este tema alcanzase celebridad. Sin embargo, aunque recibió gran cantidad de respuestas, ninguna logró sistematizar la reflexión sobre la bullshit. Actualmente ha habido intentos muy notables, como la propuesta anti-retórica de la taurascática de Fredal. Dentro de las disciplinas más propiamente filosóficas, una de las aproximaciones más interesantes al fenómeno de la bullshit es la que recientemente ha realizado Quassim Cassam desde la Epistemología de los Vicios.

La Epistemología de los Vicios funciona como reverso de la Epistemología de las Virtudes, que es un enfoque que, dentro de la Epistemología y tomando parte del vocabulario de la Ética, asume que la fuente última de los valores y normas epistémicamente relevantes está en ciertas disposiciones o competencias poseídas por el sujeto (que puede ser individual o colectivo). Consecuentemente, la Epistemología de los Vicios se centra en aquellas disposiciones o competencias que impiden de alguna manera la consecución del conocimiento. Esto puede estudiarse desde una perspectiva que entiende el vicio como defecto adquirido, que es el punto de vista del (neo)aristotelismo de Taylor y Zagszebski, o desde una perspectiva que se centre no en la adquisición o mantenimiento del defecto, si no en su carácter de obstáculo. Ese es justamente el punto de vista del obstructivismo, que es la tendencia a la que se adscribe Cassam en su *Vices of the Mind* (2019). Para Cassam un vicio epistémico es un estilo de pensamiento, un rasgo del carácter o una actitud, en todo caso reprochable, que sistemáticamente obstaculiza la obtención, compartición, gestión o retención del conocimiento. En particular, la insouciance, que podemos traducir como desinterés o despreocupación por la verdad, cae dentro de los vicios actitudinales, que son posiciones, disposiciones o perspectivas dirigidas hacia un cierto objeto. Deben ser involuntarios y gravarse de una cierta carga emotiva o afectiva. La insouciance como actitud viciosa refiere a la disposición

* pveraveg@ull.edu.es

generalmente negativa a devaluar o directamente despreciar la evidencia, los hechos o la verdad. Por ello, Cassam la relaciona directamente con la bullshit analizada por Frankfurt. Afirma que esta última es el producto de la primera, pero, dada la caracterización que ha hecho de la insouciance, esta reducción no está exenta de problemas. Mi crítica es la siguiente: tal y como Frankfurt interpretó la bullshit, la reducción de esta a insouciance puede ser insuficiente para explicarla. La despreocupación por la verdad constituye un rasgo esencial de la bullshit, pero no es el único. Gracias al contraste con los perfiles de la mentira y de la farsa, la deducción del concepto de bullshit incluye la figura de la simulación de las de intenciones del hablante. Ciertamente, la manera en que se relacionan la despreocupación respecto de la verdad y la simulación de intenciones no resulta tampoco especialmente clara en *On Bullshit*. Dos son las opciones que se oponen radicalmente: o bien predomina la simulación, y por tanto la bullshit es irreducible a insouciance; o predomina la despreocupación por la verdad, y por tanto la reducción que defiende Cassam es posible.

No obstante, aunque puede parecer que esta oposición significa una contradicción, en realidad, según creo, denota una gradualidad en la que es posible representar los diversos momentos del fenómeno de la bullshit. Además, tanto la simulación como la despreocupación son relativizables también a un cierto grado, lo que nos permite profundizar más aún en la noción de bullshit como insouciance a la vez que se establecen nuevos vicios epistémicos.

Hay un punto en el que Cassam y Frankfurt están de acuerdo y que, sin embargo, puede generar problemas a la hora de manejar el concepto de bullshit, independientemente si se es o no reducido a insouciance. Las cosas nos interesan o dejan de interesar en cierta medida. Puede haber un desinterés absoluto. Sin embargo, que este sea la norma para la definición de la insouciance es tremendamente arbitrario, aunque es lo que Cassam y Frankfurt parece que hacen. Al pensar el desinterés como una cuestión que entiende igualmente de cantidades, se puede pasar de la conceptualización de un vicio epistémico de carácter general y absoluto a uno de carácter más concreto, probablemente discursivo y muy ligado a la situación actual en la que se produce la bullshit. Si antes se postulaba una cierta tipología de la bullshit atendiendo a la cantidad de insouciance relativa a la simulación, ahora será posible diferenciar especies de bullshit dependiendo de la cantidad de insouciance que presentan.

El hecho de que la verdad no sea siempre alcanzable deberá ser igualmente tenido en cuenta en este punto. La idea de una bullshit regionalizada puede resultar interesante. Quizás, al hablar de ética, de estética o de metafísica, no pueda más que producirse bullshit. Sin embargo, caben otros enfoques para aquellos ámbitos en los que es posible sostener que la verdad simplemente no se da. Para evitar la posible caída de todos los juicios estéticos, éticos y metafísicos en la extensión de la bullshit, lo que debe hacerse es fundamentalmente tomarse las cosas en serio. Esto no salva a estas disciplinas del

sinsentido, pero sí que evita en todo caso que padezcan las actitudes epistémicamente viciosas que aquí se señalan.

Otro elemento que resulta problemático en la reducción de la bullshit a insouciance, es que, en *Vices of the Mind*, la insouciance no es una pura nada sino que también añade algo aunque sea sólo desde la dimensión afectiva antes mentada. Este añadido es el odio, la aversión, el desprecio o el asco al que se refiere el desinterés. Creo que puede resultar interesante repensar, para profundizar en la tipología de la bullshit, los modos en los que esta negatividad puede reflejarse pues no generará la misma disposición viciosa el odio por el medio de expresión, que el odio por el mensaje, por el receptor o incluso por el emisor mismo. No obstante, que la bullshit frankfurtiana cargue con la negatividad que Cassam le atribuye es discutible: el fenómeno estudiado por Frankfurt es tan amplio y tan diverso que esa limitación le resulta, según creo, completamente ajena. Aún así, que una parte de la bullshit, incluya esa negatividad puede nuevamente servirnos para generar un marco teórico completo en torno a su concepto.

Bibliografía

- [1] Battaly, H. (2015). *Virtue*. Cambridge: Polity Press.
- [2] Black, M. (1982). The Prevalence of Humbug. *Philosophic Exchange* 13 (1), 1-22.
- [3] Cohen, G. A. (2006). "Deeper into bullshit". En G. L. Hardcastle y G. A. Reich (eds.), *Bullshit and Philosophy* (pp. 117-136). Chicago y La Salle: Open Court.
- [4] Frankfurt, H. (2005). *On Bullshit*. Princeton: Princeton UP.
- [5] Fredal, J. (2011). Rhetoric and Bullshit. *College English*, 73 (3), 243-259.
- [6] Hardcastle, G. L. (2006). "The Unity of Bullshit". En G. L. Hardcastle y G. A. Reich (eds.), *Bullshit and Philosophy* (pp. 137-150). Chicago y La Salle: Open Court.
- [7] Preti, C. (2006). "A Defense of Common Sense". En G. L. Hardcastle y G. A. Reich (eds.), *Bullshit and Philosophy* (pp. 19-32). Chicago y La Salle: Open Court.
- [8] Qassim, C. (2019). *Vices of the Mind. From the Intellectual to the Political*. Oxford: Oxford University Press.
- [9] Taylor, G. (2006). *Deadly Vices*. Oxford: Clarendon Press.
- [10] Zagzebski, C. (1996). *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge*. Cambridge: Cambridge University Press.

FILOSOFÍA DEL LENGUAJE

Too much confidence in certain beliefs. Affective polarization and the possibility of error in ascribing certain mental states

Manuel Almagro*
University of Granada

Political polarisation endangers democracy. Over the last decades, scholars from different academic fields have devoted a considerable amount of effort to studying the outcomes of the current level of political division in some contemporary democracies, as well as to pinpointing its causes. It has become quite trivial to claim that increased polarisation poses a threat to the proper functioning of democracy, although this view has not always been so common. Political polarisation increases distrust in public institutions and in government (Hetherington & Rudolph 2015), makes citizens less likely to seek different perspectives on controversial topics (Valentino et al. 2008), and corrodes the proper functioning of democratic institutions (Carothers & O'Donohue 2019; Levitsky & Ziblatt 2018). Moreover, it also diminishes people's tolerance for certain issues and certain disenfranchised social groups and therefore exacerbates discrimination and violence (Mounk 2018; Wojcieszak & Garrett 2018).

The rise of polarisation is not just something that might occur as a consequence of a series of casual historical events, such as the huge development of certain technologies and the Internet, and with it the widening of our range of possibilities concerning the information we consume and the kind of people we mostly engage with (Almagro & Villanueva 2021; Sunstein 2017). Polarisation can also be triggered by the actions of political actors, used strategically to advance their agendas and obtain political gain (Cassam 2021). This strategic employment of polarisation can be carried out through different mechanisms, such as exploiting the way our informational environment works by sending personalised messages to attract independents and urge supporters to take action (Stromer-Galley 2019), or generating situations of crossed disagreement in public settings (Osorio & Villanueva 2019; Almagro, Osorio & Villanueva 2021). So, given these potential threats, we need to know as much as possible about the type of polarisation that puts democracy at risk.

I focus here on the concepts of ideological polarization and affective polarization, as they are commonly understood in literature, two prominent concepts of polarization. The former is the traditional way of conceiving polarization: two people are more polarized to

* malmagro@ugr.es

the extent that their beliefs are more distant from each other in an ideological spectrum, and a society is more polarized to the extent that the blocks in which people's opinions are clustered are more separated from each other (Bramson et al. 2017; Hetherington 2009). However, this type of polarization, based on the location of policy preferences, seems to have become useless to account for the dangers pointed out above: some democracies were experiencing serious problems of coordination and yet the polls showed that population's beliefs remained the same (Morris 2017). Affective polarization, on the other hand, is rooted in affect, and has to do with the tendency to have negative feelings toward the outgroup and positive feelings toward the ingroup (Iyengar et al. 2019). This concept seems more useful to grasp the type of polarization that endangers democracy: it was, and still is, considerably high in those societies that exhibit the problems pointed out earlier (see Boxell et al. 2020; Gidron et al. 2020; Iyengar et al. 2019). Thus, it is concluded that many contemporary democracies are not polarized in people's *beliefs*, but in their *feelings*.

Given the way that both types of polarization are commonly measured, namely through direct self-report surveys, it is my contention that one of the philosophical assumptions behind both concepts is the first-person authority thesis, at least in its weak version (Falvey 2000; Villanueva 2014): it is presupposed that participants' sincerity usually guarantees the truth of their mental self-ascription, be they beliefs or feelings. As intuitive as this thesis may be, there are growing theoretical and empirical reasons to be suspicious about their success: we are prone to fail to identify the state of mind we are in, even when we are sincere (Schwitzgebel 2008). It can be the case that the rule we sincerely say we follow is not the rule we actually follow (Wittgenstein 1953). Along with other concerns, this objection seems to undermine the ability of both concepts to measure what they try to measure.

In this paper, I hold that affective polarization provides a better account of the type of polarization that jeopardizes the well-functioning of democracy than the concept of ideological polarization because, through the feeling thermometer and the other tools employed to measure affective polarization, it is not measured just what people actually feel, but their level of confidence in the core beliefs of the political identity they identify with. To do so, I show that these tools involve evaluative language, and argue that through the evaluative use of language we usually express our attitudes, but, crucially, not any kind of attitude: those that are especially linked to action (Gibbard 2012). These attitudes provide information about our own mindset, that is, information about what can be reasonably expected from us. More specifically, they speak about people's level of confidence in certain beliefs, how impervious we are toward the arguments coming from the "other side" (Bordonaba & Villanueva 2018). In this sense, some types of affective polarization, I argue, does not have to do just with feelings toward the ingroup and the outgroup, but also with the *expression* of certain attitudes related to certain level of confidence in the core beliefs of the group people identity with.

Understood this way, certain types of affective polarization don't need to endorse the first-person authority thesis, and can avoid other objections. Moreover, this reassessment permits the concept to accommodate certain cases of polarization that otherwise would go unnoticed, such as certain polarization processes that involve a change in the degree of belief without a change in the concomitant feelings.

References

- [1] Almagro, Manuel, Javier Osorio & Neftalí Villanueva (2021). Weaponized testimonial injustice. *Las Torres de Lucca. International Journal of Political Philosophy*.
- [2] Almagro, Manuel & Neftalí Villanueva (2021). Polarización y tecnologías de la información: radicales vs. extremistas. *Dilemata, Revista Internacional de Éticas Aplicadas*, 34: 51-69.
- [3] Bordonaba, David & Neftalí Villanueva (2018). Affective polarization as impervious reasoning', *13th Conference of the Italian Society for Analytic Philosophy*.
- [4] Carothers, Thomas & Andrew O'Donohue (2019). *Democracies Divided: The Global Challenge of Political Polarization*. Washington, D.C.: Brookings Institution Press.
- [5] Falvey, Kevin (2000). The basis of first-person authority, *Philosophical Topics*, 28, 2.
- [6] Gibbard, Alan (2012). *Meaning and Normativity*. Oxford: Oxford University Press.
- [7] Hetherington, Marc (2009). Putting Polarization in Perspective, *British Journal of Political Science*, 39, 2: 413-48.
- [8] Hetherington, Marc & Thomas Rudolph (2015). *Why Washington Won't Work*. Chicago: University of Chicago Press.
- [9] Iyengar, Shanto et al. (2019). The origins and consequences of affective polarization in the United States, *Annual Review of Political Science*, 22: 129-146.
- [10] Levitsky, Steven & Daniel Ziblatt (2018). *How Democracies Die*. New York: Crown.
- [11] Osorio, Javier & Neftalí Villanueva (2019). Expressivism and Crossed Disagreements. In María José Frápolli (ed.) *Expressivisms, Knowledge and Truth*. Cambridge: Cambridge University Press, pp. 111-132.
- [12] Schwitzgebel, Eric (2008). The Unreliability of Naive Introspection, *Philosophical Review*, 117: 245-273.
- [13] Stromer-Galley, Jennifer (2019). *Presidential Campaigning in the Internet Age*. New York: Oxford University Press.

- [14] Valentino et al. (2008). Is a worried citizen a good citizen? Emotions, political information seeking, and learning via the Internet, *Political Psychology*, 29: 247-273.
- [15] Villanueva, Neftalí (2014). Know thyself: A tale of two theses and two theories, *Teorema*, 33, 3 (2014): 52-53.
- [16] Wittgenstein, L. (1953). *Philosophical Investigations*. Basil Blackwell.
- [17] Mounk, Yascha (2018). *Democracy: Why Our Freedom Is in Danger and How to Save It*. Cambridge: Harvard University Press.
- [18] Wojcieszak, M. & R. Kelly Garrett (2018). Social identity, selective exposure, and affective polarization: How priming national identity shapes attitudes toward immigrants via news selection. *Human Communication Research*, 44, 3: 247-273.

The Dual Character Concept Approach to Normative Generics

Lucien Baumgartner*
University of Zurich

Normative generics, such as the rebukes “Boys don’t cry.” or “Scientists don’t falsify data.”, are ubiquitous in everyday-language. But what exactly makes them normative?

Leslie (2015), among others, uses Dual Character Concepts (DCCs) to explain generic normativity. According to the DCC framework, the subject term “scientists”, for instance, has a dual sense: descriptively speaking, a cheating scientist is still a scientist, but in normative senses she ultimately fails to be a real or true scientist--which warrants scolding. Conversely, a painstakingly accurate accountant can be said to be “more of a (true) scientist” than our cheating fellow, even though he is not descriptively speaking a scientist. Hence, DCCs have pejorative effects on certain individuals dependent on whether they are perceived to exhibit the descriptive or normative properties. Accordingly, normative generics allow inferences to normatively shifted particular predications, e.g. “x is not a true scientist”. In that respect, I argue that embedded DCCs act similar to expressives such as “jerk” or “asshole”, even though DCCs have more specific application conditions. On a reductive reading, however, DCCs can be understood as expressives with intensionally defined quantifier-domain restrictions.

While the DCC framework works well for lexical norms in social kind-generics like the ones above, it cannot account for context-dependent normative interpretations of generics. Consider the statement “Humans are bipeds.” In a room full of joggers, this generic will pass as an inoffensive descriptive characterization. In a room with paraplegics present,-- well, not so much. How can the DCC framework explain these contextual differences in normativity?

My paper shows that the conditions for dual character normativity can be generalized from concepts to expressions (DCEs) in order to eliminate this blind-spot. For this, I examine the DCCs’ minimal felicity-conditions and contrast them with the conditions for interpreting context-dependent normative generics. By using linguistic tools from the generics literature, I show that normative generics can be analyzed as a specific variation of counterfactual definitions. With this revision, dual character normativity can explain a much larger range of phenomena, from conventionalized normative concepts to ad hoc generic expressions of norms. The popular phrase “OK boomer.”, for example, is a dismissive response often used by millennials to discredit antagonizing platitudes held by

* lucien.baumgartner@philos.uzh.ch

the baby-boomer generation. The same pejorative evaluation can be extended to millennials who hold typical boomer-beliefs, thereby mislabeling their generational affiliation. Hence, “boomer” acts similarly to “jerk”, as it expresses a negative attitude towards the target, but, in contrast to “jerk”, its use needs to be contextually licensed, because it has intensional content.

All in all, this paper represents a critical extension of the DCC framework, highlighting both its merits and limitations.

The Logic of Polyreference

Laura Delgado*
University of Lisbon

In previous work, I have defended a novel view of the semantic of proper names, i.e., the Polyreferential View. According to this view, proper names can have more than one semantic referent, indeed as many as bearers of the name there are. This view emerged partly as a response to the problem of multiple bearerhood, and it accounts for this problem in a way that preserves the referentialist's contention that proper names are fundamentally referential terms; while keeping both the intuition that names are shared by multiple bearers, and the intuition that bearing a name is the semantic relation of reference.

Polyreferentialism does not just restate that names have many different bearers but holds that each bearer of a given name *N* is a semantic referent of *N*, in an absolute (or context-invariant) sense. For a single name *N*, there are as many referential relations to individuals as bearers the name has, and the reference relation obtains between a single name and each one of its bearers individually. Thus, it is not that the name thereby refers to some plurality or loses its function as referential term and becomes instead a general term, designating a set of individuals. In other words, a polyreferential name does not have a single extension consisting of all of its bearers, but many extensions each consisting of one bearer. Thus, 'David' refers to Hume, and to Kaplan, and to Lewis, but it neither refers to Hume, Kaplan and Lewis taken together (as a plural term would) nor denotes the set of all Davids (as a predicative name would).

This has the implication that a sentence containing a polyreferential name would thereby express many propositions or contents. Given that, according to Polyreferentialism, proper names are still referential terms and what they do, being referential terms, is to contribute their referents to the propositional content of sentences in which they occur, the sentences in which they appear will also express many different contents, as many as referents for the name there are. For example, a sentence such as 'David is happy' does not express a single proposition about various Davids taken together, or about a set of Davids. Rather, it expresses many different singular propositions, each about a single David. What we said above about 'David' referring to Hume, and to Kaplan, and to Lewis, and not to Hume, Kaplan and Lewis taken together, entails that the sentence expresses the proposition that Hume is happy, and it expresses the proposition that Kaplan is happy, and it expresses the proposition that Lewis is happy. (note: The view makes a distinction between what a sentence semantically expresses (possibly many different propositions)

* lauged@gmail.com

and what a speaker that uses such sentence expresses - typically just one of the propositions the sentence semantically expresses). But it does not express the proposition that Hume, Kaplan and Lewis are happy.

This paper proposes a logic to represent this new semantic relation of polyreference and gives an account of compositionality. Given a standard way of understanding sentence truth as depending on the truth of the proposition it expresses, we have that for any sentence S , S is true when it expresses a true proposition and S is false when it expresses a false proposition. So, most sentences containing polyreferential terms will be true, and they will be false. As a result, we won't be able to reason from the truth or falsity of sentences in the standard way – e.g. 'David is happy and David is not happy' expresses many different propositions, only some of which are contradictions, for many others are true propositions, e.g., when Hume is happy and Lewis is not (among many other possibilities). More dramatically, 'David = David' is not true only (let alone necessarily true), since it does express the proposition that Hume is identical to Kaplan, among many others, and these are false propositions.

So, we won't be able to reason from the truth or falsity of sentences in the standard way. Reasoning has to live with this 'duality' so to speak: most sentences are true and they are false, although some sentences are only true, or only false (contingently or necessarily). A conjunction of two atomic sentences that are both true and false would seem also to be true and false. Likewise for other complex sentences. Somehow this duality in value is preserved. But it seems that if the premises are both true and false, then at least one proposition expressed is true, and so the conclusion has got to be true (even if also false) on account of that proposition being true. Another way to put it is that the truth of the premises ensures that the conclusion is not going to be only false (for it would be both true and false).

Maybe this is good enough for the purpose of reasoning with names. The thought above is in line with Priest's Logic of Paradox (LP). Priest also offered a sub-logic of LP to handle multiply denoting terms (Priest, 1995). I use this logic to represent polyreference and to handle the multiple propositionalism it entails. Although in this framework, some standard reasonings would not be valid, the paper argues that this is not necessarily a bad result. I suggest that when taking the fact that names have multiple bearers seriously, our intuitions about the truth and falsity of sentences, and of the validity of inferences may change. The logic proposed, however, can cope with most of our ordinary reasonings with proper names.

References

- [1] Priest, G. (1995). 'Multiple Denotation, Ambiguity, and the Strange Case of the Missing Amoeba'. *Logique Et Analyse* 38: 361-73.

Provocative insinuations

Álvaro Domínguez Armas*
IFILNOVA

Andrés Soria Ruiz**
IFILNOVA

1 Introduction

In this paper we are concerned with utterances that, without explicitly constituting hate speech, nevertheless convey a hateful message (Maitra & McGowan, 2012; McGowan, 2019):

[1] Iraqi Refugee is convicted in Germany of raping and murdering teenage girl.
~> Iraqi refugees are despicable.

[2] Un marroquí, detenido por esconder 187 kilos de hachís bajo la cuna de su hijo.
A Moroccan, arrested for hiding 187kg of hashish under his son's cradle.
~> Moroccan people are despicable.

The choice of 'Iraqi refugee' and 'marroquí' does not seem arbitrary. To the contrary, it is responsible for inviting an inference against Iraqi refugees and Moroccan people. Call this a racist inference (RI). This can be seen by noting that, if we substitute 'marroquí' in [2] for another predicate, such as 'postman', the inference disappears:

[3] A postman, arrested for hiding 187kg of hashish under his son's cradle.
??~> postmen are despicable.

We begin by disregarding the immediate hypothesis that this inference is a consequence of the presence of a negative predicate: note that an inference against the referred group appears even with positive examples (even if the inference is weaker):

[4] Cuatro vecinos gitanos evitan que una mujer sufra una violación en plena calle
Four Romani neighbours saved a woman from being raped in the middle of the street.
~> Contrary to expectations, these 4 Romani people are not despicable.

* alvarodguez@hotmail.com

** andressoriaruiz@gmail.com

In this case, the headline seems to be remarking on the weirdness of the situation, where the saviours, and not the attackers, belong to the Romani community. Thereby, the racist inference against the Romani community is still produced, not via a negative predicate, rather, by reporting on an atypical situation.

We defend that RIs cannot be described as *slurs* (§2), *ethnic and social terms used as insults* (ESTIs, §3) or *conversational implicatures* (§4). We propose that RIs are a type of *insinuation* (Camp, 2018), in particular, a *provocative* insinuation (§5).

2 RIs as slurs

Characterizing ‘Iraqi refugee’/‘marroquí’ in [1]-[2] as a slur might be suggestive insofar as *slurs* give rise to a similar inference (Hess, 2020a, 2020b):

[5] Bill is a N*
~> black people are despicable.

However, there are two important differences between our phenomenon and slurs. First, slurs are lexical items which are conventionally derogative. By contrast, ‘marroquí’ is not conventionally derogative; in fact it is the neutral counterpart of a Spanish slur (‘moro’). Secondly, the projective profile of slurs and RIs is different, as illustrated by the contrast between [5] and [6]. As it’s well established, the derogatory content of slurs projects out of environments such as negations, questions or modals.

[6] a. Bill is not a N*
b. Is Bill a N*?
c. Bill might be a N*
~> black people are despicable.

By contrast, PIs do not project in the same way:

[7] a. Iraqi refugee is not convicted of raping and murdering a teenage girl.
b. Is an Iraqi refugee convicted of raping and murdering a teenage girl?
c. Iraqi refugee might be convicted of raping and murdering a teenage girl.
?? ~> Iraqi refugees are despicable.

Thus, we conclude that RIs are a different phenomenon from slurs.

3 RIs as ESTIs

In some environments, an otherwise neutral ethnic or social term can adopt a pejorative meaning (Castroviejo et al., 2020):

[8] Juan es un gitano.

Juan is a roma

[8] attributes to Juan a set of stereotyped—negative—properties of Romani people. In our examples, we also find ethnic or social terms triggering stereotypes.

However, there is a crucial difference between an ESTI such as [8] and the RIs in [1]/[2]: [8] does not entail that Juan is Romani but [1]/[2] entail that the subject is an Iraqi refugee / Moroccan. We take this contrast to show decisively that RIs are not ESTIs.

4 RIs as conversational implicature

Might RIs be a kind of *conversational implicature* (CI, Grice, 1989)? In a CI, the hearer makes an inference based on a contribution that violates a conversational maxim, while assuming that the speaker is cooperative. Here is a standard example, based on a violation of the Maxim of Relation (Grice, 1989, p. 32): Imagine a subject a, whose car is immobilized, approached by another subject b. The following communicative exchange takes place:

[9] a: I am out of petrol.

b: There is a garage round the corner.

Here, a might infer that b is mentioning the garage because it is open and has petrol to sell.

Similarly, the line of reasoning from [1] to the RI might be based on the idea that, if the speaker mentioned the person's nationality or origin, it must be relevant, thereby generating the RI.

However, RIs and CIs differ in the following way: when a CI is *cancelled*, the speaker appears *uncooperative*. For instance, if a in [9] replies 'Do they sell petrol there?' and b answers 'No, they don't!', b will appear uncooperative. The reason for this is that, without the implicature, b's original contribution seems completely irrelevant to the conversation. This does not happen with [1]/[2]. Imagine that the speaker is accused of racism by uttering [1]. By replying 'I didn't

say/mean that!' or 'Iraqi refugees are wonderful people' the speaker does not come across as uncooperative. This suggests that [1]/[2] do not violate any conversational maxim, and thus that the RI is not rescuing an otherwise uncooperative contribution (see Cohen & Kehler 2021).

5 RIs as insinuations

We defend that RIs can be characterized as *insinuations* (Camp, 2018; Fricker, 2012).

Insinuations are 'off-record' pragmatic inferences triggered by certain sentences. If someone utters

[10] You know that Obama's middle name is Hussein. I'm just saying

the hearer might infer, e.g., that Obama supports radical Islamism (Camp, 2018, p. 43).

Insinuations are *deniable* and *disavowable*: the speaker can deny their content as well as having intended to convey it without appearing uncooperative. For instance, if the speaker of [10] is confronted with: 'Are you suggesting that Barack Obama is a radical Islamist?', they can answer 'Of course not!'. Similarly, if the speaker of [1] is confronted with 'Are you suggesting that Iraqi refugees are despicable?' The speaker can answer 'Of course not!'

We finish by arguing that insinuations like [1]-[2] are *provocative* because they place hearers in a frustrating position, as they are forced to *succumb* to their *provocative* dimension. Possible responses can be categorized as follows:

- 1) *Ignore the RI*. This strategy risks that the RI might be incorporated into the conversational *common ground*.
- 2) *Confront the RI*. There are two risks that this strategy poses. First, the RI might become the topic of the conversation. Participants abandon the flow of the dialogue to discuss whether or not there is a content hidden under the utterance. Secondly, the content of the PI might gain credibility simply by coming under discussion.

In sum, any response that targets the RI seems *inefficient* at blocking the inference and risks turning the conversation into a racist one, thereby serving the purposes of the insinuator.

References

- [1] Camp, E. (2018). Insinuation, Common Ground, and the Conversational Record. In D. Fogal, D. W. Harris, & M. Moss (Eds.), *New Work on Speech Acts* (pp. 40–66). Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780198738831.001.0001>
- [2] Castroviejo, E., Fraser, K., & Vicente, A. (2020). More on pejorative language: insults that go beyond their extension. *Synthese*.
- [3] Cohen, J., & Kehler, A. (2021). Conversational eliciture. Forthcoming in *Philosopher's Imprint*.
- [4] Fricker, E. (2012). Stating and Insinuating. *Proceedings of the Aristotelian Society Supplementary Volume*, 86(1), 61–94.
- [5] Grice, P. (1989). *Studies in the Way of Words*. Harvard University Press.
- [6] Hess, L. (2020a). Practices of slur use. *Grazer Philosophische Studien*, 97(1), 86–105.
- [7] Hess, L. (2020b). Slurs and Expressive Commitments. *Acta Analytica*, July.

- [8] Maitra, I., & McGowan, M. K. (2012). Speech and Harm: Controversies Over Free Speech. In *Speech and Harm: Controversies Over Free Speech*. Oxford University Press.
- [9] McGowan, M. K. (2019). *Just Words: On Speech and Hidden Harm*. Oxford University Press.

The Classificatory Model of Communication and Indexical Utterances

Ainhoa Fernández*
University of the Basque Country

In the philosophical discussion on communication there is a perspective which enjoys considerable, if not majority, acceptance. Let us call it *The Transmission Model of Communication*. According to this model, communication consists in the transmission or replication of propositions, so when a speaker makes an utterance the hearer reconstructs the proposition that the speaker intends to communicate. The strongest version of this thesis is the Naïve conception of communication (Heck 2002) - sometimes called the belief transfer model (Egan 2007) or the FedEx model (Weber 2013). According to the transmission model there are three elements which intervene in a communicative process: (a) the propositional content of the speaker's thought, (b) the proposition grasped by the hearer, and (c) the proposition literally expressed by the sentence. In the Naïve conception the three elements are actually the same because the underlying idea is that the content, the information which is transmitted, undergoes no change.

There is a well-known problem presented by the Naïve conception: to explain cases in which we try to communicate indexical thoughts. Indexical thoughts are restricted by the context and essentially linked to it. The problem with communication of indexical utterances is that there is a part of the content which can only be entertained by the speaker and not by the hearer, so that it cannot be transmitted or replicated. Perry (1979) called this phenomenon limited accessibility. All the different versions of the transmission model aim to offer a solution to the problem that indexical utterances pose for the naïve theory. On the one hand, we have approaches in which a transformation-recentering process is carried out and in which no content is required to be shared by the speaker and the hearer (Gibbard 2012, Weber 2013). On the other hand, there are those of a Fregean substitutive spirit, according to which what is communicated is a substitutive (ersatz) thought or proposition in such a way that this new thought is shared by the speaker and the hearer (Kölbel 2013, Torre 2010).

What I will provide in this paper is an alternative to the transmission model of communication: the Classificatory Model of Communication. This model is grounded in the act-type view of propositional content (Hanks 2011, 2015; Soames 2015). I shall argue that in this new conception cases of indexical communication are not a problem and that,

* afdezsoutullo@gmail.com

therefore, the model does not need any adjustment or modification (substitutive, transforming or recentering process) to explain how successful communication is achieved in such cases, which simply fit smoothly into the model. Additionally, I will contend that communication in the act-type theory is to be understood as description and explanation of action.

In the first place, I will argue that the reason why all the variants of the transmission model are forced to carry out a process of modification of the utterance performed by the speaker is that they all maintain a Fregean conception of content. More particularly, I will argue that all these theories work under what I will dub the uniqueness of content assumption: the idea that there is one and only one proposition that can be considered as the content of an act of assertion.

According to the act-type view of propositional content propositions are types of actions which function as classifying devices for describing, identifying, and individuating our thoughts and speech acts. They are also the derivative bearers of truth conditions, which are obtained through instantiated or possible tokens of each type. Thus, when a speaker makes a token utterance the hearer classifies it under a certain type. On the classificatory model, successful communication consists in a hearer classifying the speaker's utterance under an appropriate type. Among these types we find that some are more specific than others in the sense that some contain more information and details than others, that is, some are more coarsely-grained and others which are fine-grained. When we classify an utterance there is not a single type which is the only correct one, but there is a variety of types under which we can classify it. There is no assumption of uniqueness built into the act-type theory. Actually, it is just the opposite, any token act falls under many types.

In the act-based view is defended that there are semantic reference types which are certain types of reference acts, for example, a type of reference act I perform when I use the name 'Antonioni' or the first person pronoun 'I'. Accordingly, "two acts of reference fall under the same semantic reference type just in case anyone who is semantically competent with the terms used in those acts will know that those two acts corefer" (Hanks 2015, 8). This feature combined with the variety of types is what will allow us to accommodate the indexical cases without major trouble.

As I will argue, indexical utterances do not pose a problem for the classificatory model. Suppose a speaker says "I am extremely tired". On the classificatory approach, the speaker has performed a type of action that only she can perform. In that sense, the content of her utterance is of limited accessibility. However, there is no barrier to a hearer classifying this utterance under that type, even if the hearer is not capable herself of performing a token of that type. In this way, indexical utterances can be easily integrated into the classificatory model of communication.

References

- [1] Egan, A. (2007) Epistemic Modals, Relativism and Assertion. *Philosophical Studies* 13: 31–22.
- [2] Gibbard, A. (2012). *Meaning and Normativity*. Oxford: Oxford University Press.
- [3] Hanks, P. (2011) Structured Propositions as Types. *Mind* 120 (477): 1-52.
- [4] Hanks, P. (2015) *Propositional content*. Oxford: Oxford University Press
- [5] Heck, R. (2002) Do Demonstratives Have Senses? *Philosopher's Imprint* 2: 1-33.
- [6] Kölbel, M. (2013) The Conversational Role of Centred Content. *Inquiry*. 56: 97-121.
- [7] Perry, J. (1979) The Problem of the Essential Indexical. *Noûs* 13: 3-21
- [8] Soames, S. (2015). *Rethinking language, mind, and meaning*. Princeton: Princeton University Press.
- [9] Torre, S. (2010) Centred Assertion. *Philosophical Studies* 150: 97–114.
- [10] Weber, C. (2013) Centred Communication. *Philosophical Studies* 156: 205–23.

Singular Thought and Strong Acquaintance

Heimir Geirsson*
Iowa State University

There is an important cognitive difference between a thought that is directed towards a particular object and a thought that is not so directed but is instead about a certain kind of objects. For example, there is a difference between my thoughts about my brother and my thoughts about brothers more generally. The former is an example of a singular thought while the latter exemplifies general thoughts.

Most direct reference theorists adhere to what I will call the orthodox view when it comes to acquaintance, namely the view that one can be sufficiently acquainted with an object and so obtain a singular thought about it in virtue of being at the receiving end of a use of a name of the object that stretches back to an initial baptism of it.

I will argue that we have good reasons to doubt the truth of the orthodox view and favor instead an account that advocates a strong acquaintance requirement.

The direct reference theory provides the backbone of the orthodox view. According to the direct reference theory a reference of a name is not determined by how an object fits a given set of descriptions. Instead, reference is secured via a causal chain, where one user of a name passes it on to another who then intends to use it with the same reference. The reference of the name is therefore not satisfactorial. Because the reference of the name is not satisfactorial the advocates of the orthodox view have been quick to conclude that the thought that results from sincerely assenting to a simple sentence that contains the name of an object, and the thought that one reports with an embedded simple sentence containing the name, is a singular thought. Being on the receiving end of the causal chain of a name, according to the orthodox view, is acquaintance enough for having a singular thought about the relevant object.

Robin Jeshion puts the points as follows: “If you finally met me, would you thereby better understand the term ‘Robin Jeshion’? Surely this is something that the Millian denies.” Again, the point is that according to the advocates of the orthodox view there is no “additional meaning” beyond the referent to be found in names and so there is no “additional understanding” to be had once one has acquired the name. Acquiring a name of an object enables one to have singular thoughts about it. Jeshion seems to be echoing a point made earlier by Kent Bach when he argues that when a name is passed on “a speaker cannot just express but can actually display his *de re* way of thinking of the object and

* geirsson@iastate.edu

thereby enable the hearer to think of it in the same way.” The underlying assumption that Bach appears to be working with is the following:

The testimony requirement: A sufficient condition for one having a singular thought of an object is that one acquire a name of the object, the name having being initially introduced with an acquaintance relation.

Note that the requirement allows one to be on the receiving end of a long chain of use, stretching back to an initial baptism, and still have singular thought about the object named.

Gareth Evans provides an example that clearly points out that there is a problem with Bach’s account of displaying or inheriting a way of thinking via the use of names as well as Jeshion’s claim about understanding names. Suppose that person X joins a group that is talking about a certain Louis. X listens in for a while and then joins in the conversation with appropriate uses of the name ‘Louis’. It certainly seems that he is, when doing so, successful in referring to the same Louis that his friends are talking about. The discussion is about King Louis XIII. If that is so then Jeshion, as well as most Millians, are committed to attributing to X singular thoughts about Louis XIII. Suppose now that due to some massive errors X comes away from the discussion believing that Louis is a basketball player. Evans comments on this, writing that it is “outrageous to say that he believes that Louis XIII is a basketball player. The notion of (singular thought) has simply been severed from all the connections that made it of interest.”

It appears to me that Evans is pulling on the right intuitions here. Even though the subject in his story comes away using the name ‘Louis’, we should not claim that the speakers displayed their way of thinking about Louis and passed it on to our subject, nor should we say that he has singular thoughts about him. Even though the name is passed on, singular thought is not.

There is a way, I will argue, to explain the intuitions that drive Evans’ example while accommodating the direct reference view, but The Testimony Requirement falls by the wayside as a result.

Several philosophers have suggested that causal connections other than the one required by testimony is sufficient for one having singular thoughts about objects. For example, Jeshion and Recanati allow that Leverrier have singular thoughts about Neptune without ever perceiving the planet. It suffices, on their account, that he has seen the appropriate causal traces of Neptune, namely the perturbations in the orbit of nearby known planets and that, of course, he satisfies Jeshion’s significance condition and Recanati’s requirement that he later become more directly acquainted with it. Relying on causal connections of the kind described above is not likely to be helpful in clarifying the nature of singular thought as they are too permissive, making it too easy to acquire singular thoughts. Instead I argue that we need a strong acquaintance relation in order to have singular thoughts of ordinary objects. This, I will argue, is in tune with Russell’s initial introduction of singular thoughts, as knowledge by acquaintance, in 1903.

El problema de la dimensión ontológica en la hipótesis de la relatividad lingüística

Carlota G. Llorente*
Complutense University of Madrid

En 2011, García Suárez ya advertía de los problemas que genera asumir versiones radicales de la hipótesis de la relatividad lingüística. De un modo informal, y siguiendo a Blanco Salgueiro (2017), la hipótesis de la relatividad lingüística, también conocida como *Hipótesis Sapir-Whorf*, puede definirse como la idea de que las lenguas difieren entre sí de modo no trivial, y estas diferencias impactan de una manera u otra en el pensamiento. Así, se generarían diferencias cognitivas entre hablantes de distintos idiomas¹. La intensidad o alcance del impacto es lo que diferencia las tendencias más drásticas de las más débiles. Teniendo en cuenta estas diferencias de grado, se puede establecer una división entre lo que Blanco Salgueiro (2017) define como influencismo, de carácter más liviano, frente a determinismo lingüístico.

García Suárez (2011), aunque sin dejar de contemplar hipótesis neowhorfianas más sutiles, se centra en dirigir la crítica contra el determinismo lingüístico acusado atribuido a Whorf (2011), a saber, la tesis de que el lenguaje determina fuertemente el pensamiento. De las diferentes críticas que plantea en el artículo, la que nos atañe aquí tiene que ver con el problema de la calibración, que el autor enuncia de la siguiente manera:

Si solo podemos acceder al mundo a través de las categorías que nos impone nuestra lengua y resulta que lenguas diversas imponen tiránicamente concepciones del mundo inconmensurables, ¿cómo podríamos librarnos de esas cadenas para ver y segmentar el mundo de modo lingüísticamente inocente? ¿Dónde podríamos encontrar puntos fijos para el calibrado? (García Suárez, 2011, p.70)

Aunque en este pasaje García Suárez se centra en el problema de la inconmensurabilidad, sumado a la tesis determinista, más adelante no pierde la oportunidad de señalar los deslices ontológicos que cometen tanto Sapir como Whorf.

* carlotaglg37@hotmail.com

¹ Cabe mencionar que, como bien señala Blanco Salgueiro (2017), realmente la hipótesis de la relatividad lingüística es el resultado de la asunción de dos premisas diferentes. Por un lado, la diversidad lingüística, esto es, que las lenguas difieren entre sí de manera relevante. Por otro lado, el impacto cognitivo del lenguaje, que asume que el lenguaje, independientemente de la lengua hablada, afecta de manera no trivial al pensamiento. Son las distintas combinaciones que se dan entre estas dos tesis y sus variantes las que dan como resultado numerosas versiones de la hipótesis de la relatividad lingüística.

El problema va más allá. Para poder plantear una posible solución al problema del calibrado primero es necesario tener una noción clara de cómo se relaciona, no solo el lenguaje con el pensamiento, sino con el mundo. Muchos autores partidarios de la relatividad lingüística han trabajado en desentrañar las conexiones que pueden existir entre las diversas lenguas y la cognición de los individuos, pero han olvidado tener en cuenta el factor ontológico: el mundo.

Es posible deducir que, si de alguna manera la lengua afecta al pensamiento, también afectará a la manera en que se concibe el mundo. Entonces, si asumimos que la realidad solo puede ser conceptualizada a través del lenguaje y diferentes lenguas imponen visiones diferentes de la realidad, ¿cómo es posible establecer parámetros objetivos de comunicación o relaciones intersubjetivas entre hablantes de diferentes idiomas? Parece que, si asumimos con Sapir (1949) y Whorf (2011) que existe una segmentación semántica de la realidad y hay una fuerte determinación del pensamiento por cada lengua, no existirá un único mundo en el que todos habitemos, sino que hablantes de lenguas estructuralmente distinta habitarán mundos diferentes. Si estamos encerrados en la jaula de nuestro lenguaje y a su vez, cada lenguaje es un mundo independiente del resto, nuestras visiones del mundo diferirán radicalmente de un hablante a otro, afectando incluso a lo que se considera admisible en la ontología.

Esto puede resultar altamente contraintuitivo, sobre todo si se mira desde una perspectiva materialista dura. Con precaución de no abusar de la obscura noción de sentido común, el hecho de que podamos viajar y visitar diferentes lugares del mundo de manera exitosa, comunicándonos felizmente con hablantes de otros idiomas, hace de la defensa de la relatividad ontológica radical una tarea un tanto extravagante. Da la impresión de que, si viviésemos en mundos absolutamente distintos, esto no podría darse. Nuestras acciones, tanto conductuales como lingüísticas, no se realizarían con éxito.

Pero el defecto no se agota en la defensa de una relatividad ontológica poco plausible. Otro inconveniente principal, que se suma a la poca conveniencia de la tesis mantenidas, es que, al plantear sus teorías, los autores pro-relatividad acostumbran a hacer juicios que establecen relaciones con la ontología incompatibles entre sí. Esto desemboca inevitablemente en inconsistencias que hacen perder plausibilidad a las hipótesis planteadas.

El caso de Whorf es uno de los más notables: en unas ocasiones admite que el mundo es un flujo caleidoscópico de impresiones organizadas lingüísticamente en nuestra mente y en otras que los hechos son diferentes para hablantes de lenguas estructuralmente distintas. Parece así que en el primer caso existiría una realidad común, similar a un noúmeno kantiano, al que accedemos a través de nuestras categorías lingüísticas. En cambio, en los pasajes como el segundo, la insinuación es de corte más radical: lenguas distintas, mundos distintos.

Dada la acentuada despreocupación por las cuestiones ontológicas que existe entre los autores clásicos de la relatividad lingüística, tal y como señala García Suárez (2011), y poniendo énfasis en que ni siquiera este interés ha conseguido despertarse en los actuales

autores neowhorfianos, se genera una necesidad imperante. Es tarea del filósofo clarificar las relaciones existentes entre la tríada lenguaje-pensamiento-mundo, estableciendo los nexos pertinentes y señalando las consecuencias ontológicas que se derivan de las múltiples variantes que adopta la hipótesis de la relatividad lingüística hoy en día².

Quizá uno de los factores que haya influido en este descuido guarde relación con la escasa formación filosófica de muchos de los adalides de esta hipótesis. Como bien es conocido, Whorf era un ingeniero metido a lingüista a tiempo parcial y Sapir antropólogo y lingüista. Las nuevas remesas, aunque ligeramente más prometedoras, siguen brillando por su ausencia de formación filosófica: Lucy (2016), Levinson (Evans & Levinson, 2009; Levinson, 2001), Boroditsky (2001), Everett (2012), Slobin (2001), entre otros. Todos ellos, grandes representantes actuales que trabajan en la plausibilidad de la hipótesis, provienen del campo de las ciencias sociales.

Con esto no se pretende otorgar ningún privilegio a la empresa filosófica frente a otras disciplinas, sino simplemente señalar una de las posibles causas del despiste ontológico de la hipótesis. El filósofo, por la naturaleza de su formación, no puede evitar detectar inconsistencias. Cuando se detectan afirmaciones, quizá ligeramente proferidas, que implican tesis incompatibles sobre la manera de relacionarnos con el mundo, surge la necesidad de emprender un análisis del asunto. Y esclarecer la dimensión ontológica de estas hipótesis reforzará la potencia explicativa de las mismas.

La cuestión de si los hablantes de diferentes idiomas habitan un mismo mundo pero con diferentes etiquetas, mundos completamente distintos en un sentido ontológico fuerte o si las diferencias lingüísticas no afectan siquiera a nuestros compromisos ontológicos, es difícil de dirimir. Sin embargo, los retos conceptuales nunca han asustado a los filósofos.

Bibliografía

- [1] Blanco Salgueiro, A. (2017). *La relatividad lingüística: (variaciones filosóficas)*. Akal.
- [2] Boroditsky, L. (2001). Does Language Shape Thought?: Mandarin and English Speakers' Conceptions of Time. *Cognitive Psychology*, 43(1), 1-22. <https://doi.org/10.1006/cogp.2001.0748>
- [3] Evans, N., & Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences*, 32(5), 429-448. <https://doi.org/10.1017/S0140525X0999094X>
- [4] Everett, D. (2012). *Language: The Cultural Tool*.

² Quine (1969) dedicó buena parte de su obra a la relatividad ontológica en estrecha conexión con el lenguaje. Es un autor que puede dar grandes pistas de los caminos a considerar.

- [5] García Suárez, A. (2011). La relatividad lingüística clásica y sus avatares actuales: razones para el escepticismo. *Revista Española de Lingüística*, 41(2), 57-94.
- [6] Language Acquisition and Conceptual Development. (2001). In M. Bowerman & S. Levinson (Eds.), *Language Acquisition and Conceptual Development*. Cambridge University Press. <https://doi.org/10.1017/cbo9780511620669>
- [7] Levinson, S. C. (2001). Covariation between spatial language and cognition, and its implications for language learning. In *Language Acquisition and Conceptual Development* (pp. 566- 588). Cambridge University Press. <https://doi.org/10.1017/CBO9780511620669.021>
- [8] Lucy, J. A. (2016). Recent Advances in the Study of Linguistic Relativity in Historical Context: A Critical Assessment. *Language Learning*, 66(3), 487-515. <https://doi.org/10.1111/lang.12195>
- [9] Lucy, J. A., & Gaskins, S. (2001). Grammatical categories and the development of classification preferences: a comparative approach. In *Language Acquisition and Conceptual Development* (pp. 257-283). Cambridge University Press. <https://doi.org/10.1017/CBO9780511620669.011>
- [10] Quine, W. V. (1969). Ontological Relativity. In *Ontological Relativity and Other Essays* (pp. 26- 68). Columbia University Press. <https://doi.org/10.7312/quin92204-003>
- [11] Sapir, E. (1949). *Language*. Harcourt, Brace & Co.
- [12] Slobin, D. I. (2001). Form–function relations: how do children find out what they are? In *Language Acquisition and Conceptual Development* (pp. 406-449). Cambridge University Press. <https://doi.org/10.1017/CBO9780511620669.016>
- [13] Whorf, B. L. (2011). *Language, thought, and reality : selected writings of Benjamin Lee Whorf* (J. B. (John B. Carroll (ed.)). Martino Publishing.

Grice and Davis on Meaning

Josep Macià*
Universitat de Barcelona

I Grice's analysis

Paul Grice (Grice 1957, 1968, 1969), offered a very influential account of what it takes for an individual to mean something on some given occasion. On Grice's view, to mean that p to an audience A is to perform some action with some complex intentions regarding p and A . The classic Gricean account of speaker's meaning might be formulated as in (1)

- (1) U means_{NN} that p by 'uttering' X if and only if U utters X with the intention
- (a) that the audience A believes that p
 - (b) that A recognizes that U has the intention in (a)
 - (c) that A 's recognition that U intends (a) to be a reason for A to believe that p

I think there are good reasons for thinking that this analysis is on the right track, even if it is known that it faces a number of significant objections. Two of these objections are:

1. This analysis cannot explain why very young children that we have good reasons to believe are not able to attribute beliefs to others (let alone higher order mental states, like beliefs about intentions about beliefs) are, nevertheless, able to mean (when a 3-year old utters "I want pizza. Do not want broccoli", she does mean that she wants pizza and does not want broccoli). Gluer and Pagin (2003) make a similar point regarding some autistic subjects, who are unable to pass the various false-belief tasks, but who, on the other hand, have linguistic abilities that it seems must qualify them as speakers.
2. Audience-less cases: When rehearsing a speech, I might utter sentences and mean what they mean, even if there is no audience, and the clauses in the Gricean analysis can therefore not be satisfied. Or I might talk to a baby, or to my dog, or even to a plant, though in those cases I could not rationally intend my "audience" to form any of the complex beliefs that the Gricean analysis would require.

II Davis' alternative account

Wayne Davis (Davis 2003) offered a refreshingly simple alternative account to the Gricean classical analysis of speaker meaning. Davis' account has been claimed to avoid

* josep.macia@ub.edu

in a simple and elegant manner the two kinds of problems that I just presented for the Gricean analysis. In my paper, I argue that Davis' account does not provide a viable alternative to the Gricean analysis because (i) it is subject to new difficulties of its own, (ii) it has a complexity similar to that of the Gricean analysis, (iii) the Gricean analysis can be modified so as to successfully account for those cases that were taken to favour Davis' account.

For lack of space, here I will just comment on (ii) and, still, very briefly:

III Is Davis account substantially simpler than Grice's?

Davis' account is based on the two principles:

- (2) **Meaning:** *S* meant that *p* iff *S* either directly or indirectly expressed the belief that *p*.
- (3) **Expression:** *S* expressed the belief that *p* iff *S* performed a publicly observable action *e* as an undisguised indication that *S* has the occurrent belief that *p*.

Several notions in these clauses would require, of course, clarification. But for present purposes, let's just say that *e* is an indication of an event, if *e* is evidence that the event took place. And to say that *e* is an *undisguised* indication that *S* has a belief that *p* is to say that in producing *e*, *S* is not pretending to unintentionally manifest the belief that *p* (as would happen if, say, I fake a smile intending my "audience" to take it for a genuine smile and to think that I am happy).

But now notice: there are cases in which by producing a fake (non-spontaneous) smile an utterer comes to mean something. Say, A asks me: Are you happy? And I react by looking at A and intentionally producing a smile for 1 second (so, a non-natural, fake, smile), in order to communicate that, yes, I am happy. By producing this fake (non-spontaneous) smile, I meant that I am happy. For Davis' account to be able to accommodate this case as a case of meaning, my fake smile must be an undisguised indication that I (believe that I) am happy.

But given that A knows that my smile is fake, what would bring A to believe that I am happy on the basis of my smile?

Well, it seems that just the fact of realizing why I produced this fake smile. That is, realizing that I want her to believe that I am happy.

Therefore, in order for cases like this one to meet Davis' condition (3) it seem that it is actually necessary that they meet also a condition with the same kind of complexity as Grice's condition (c).

References

- [1] Davis, W. A. *Meaning, Expression, and Thought*. New York: Cambridge University Press, 2003.
- [2] Glüer, K. & Pagin, P. (2003). "Meaning Theory and Autistic Speakers". *Mind and Language* 18 (1):23–51
- [3] Grice, H. P. *Studies in the Way of Words*. Cambridge, MA: Harvard University Press, 1989.

The effect of slurs: World-ordering and the impossibility of retraction

Alba Moreno Zurita*
University of Granada

The aim of this paper is twofold. First, we analyze the proposal presented by Marques & García-Carpintero (2020) to account for the meaning of slurs and offer a different way of modelling the effect that slurs have on the common ground. Second, we show that a presuppositional account like the one advanced by these authors is the appropriate one when it comes to explain the difficulties faced by speakers who try to retract a slur.

According to Marques & García-Carpintero, when someone utters a slur, it is necessary for the utterance to be felicitous that a reactive attitude (see Kenny 1963) against the target group is already presupposed in the common ground. This is what *expressive presuppositions* consist in: they function like propositional presuppositions because they are what is necessary for a certain utterance to be felicitous, but the requirement for this is that there is a certain reactive attitude, and not a proposition, in the common ground. When we utter a slur and the target group is not yet within the extension of the reactive attitude of contempt, accommodation occurs so that the utterance is felicitous.

It seems to us that this proposal has two aspects that make it stand out among other presuppositional accounts. First, they flesh out what it means for content to be expressive. By doing so in a nonpropositional way, they account for slurs' special connection with action—in particular, for their derogatory character, which is what makes them dangerous. Second, by situating this content at the presuppositional level, they account for its elusive character. A presuppositional treatment, as we will see later, can thus explain the intuition that it is very complicated to retract the utterance of a slur.

However, we believe that the way these authors model the effect that slurs have on the common ground can give rise to a psychologistic interpretation (Ayala & Vasilyeva 2015). As we have said, a sentence has an expressive presupposition whenever its utterances require a particular reactive attitude to be part of the common ground to be felicitous. A way of understanding what it means for this attitude to be part of the common ground is to take speakers to share it, at least for the purposes of the conversation. This is what drives us to characterize this proposal as psychologistic, since it makes the derogatory effect of slurs depend on the particular attitudes of speakers. Once the derogatory effect of slurs is characterized in this way, we might be inclined to think that situations of injustice stem from implicit attitudes and make intervention measures depend on them too.

* almorenozurita@gmail.com

To mend this, we propose to understand slurs as ordering the possible worlds in the common ground in a certain way; in particular, in a way that corresponds to certain preferences of the speaker, as proposed by Stanley (2014: 144). Stanley relies on the distinction between at-issue content and not-at-issue content to explain how certain terms that we find in political discourse have an effect that directly alludes to the audience's emotions. At-issue content is that understood as proposed to be added to the common ground, while not-at-issue content is directly added to the common ground, without any need for the audience to accept it (Stanley 2015: 135). Not-at-issue content is added to the common ground without us having the chance to decide whether we accept it or not. According to Stanley, propagandistic discourse features not-at-issue content whose effect can be modeled along the lines of Starr's (2020) account of imperatives. Starr models the effect that imperatives have on the common ground as that of ordering the possible worlds in it in accordance with their preferability. For instance, if Víctor commands his child Pablo not to eat any more cookies, the worlds in the common ground are ordered so that those in which Pablo stops eating cookies are ranked above those in which he does not.

Our proposal can be seen as a combination of the expressive presuppositional treatment applied by Marques & García-Carpintero (2020) and the ordering of worlds in the common ground proposed by Stanley (2015) in regard to derogatory discourse. The requirement we propose for an utterance containing a slur to be felicitous is that there be a certain ordering of worlds in the common ground; and if there is not, it is accommodated. Relying on world-ordering allows us to explain, from a structural perspective, the derogatory character of slurs. When the world-ordering takes place, the socio-normative space inhabited by speakers is also modified in such a way that people who belong to the target group are left in a disadvantageous position vis-à-vis the one who utters the slur.

As advanced above, moreover, we think that a proposal along these lines is in the best position to explain why slurs are so hard to retract (Marques 2018). We may want to find ways in language that allow us to be the kind of people we want to be, and retraction is usually useful for this. However, it seems that, when we try to retract an utterance featuring a slur, we cannot undo the damage we have caused –we cannot “go back in time” to the moment before we pronounced the slur. The hypothesis we want to test in this paper is that presuppositional accounts of slurs, such as the one defended above, allow us to explain why this is so. When I say “I take that back”, what I am taking back is the asserted content, not the presupposed content. The expression “I take that back” is elliptical for “I take what I said back”, not for “I take what I presupposed back”. To see this, consider that, while we can find actual occurrences of “I take what I said back” in ordinary conversations, “I take what I presupposed back” is a philosophical construct. If this is correct, it would mean that we cannot retract the presupposed content –if we say “I take that back” trying to retract a presupposition, all we will do is retract the assertion that triggered that proposition. Now, if the derogatory content of slurs is presupposed, this would explain why we cannot undo the damage associated with the utterance of a slur– because its

derogatory content cannot be targeted by an utterance of “I take that back”, because “I take that back” targets asserted content and derogatory content is presupposed.

References

- [1] Ayala, S. & Vasilyeva, N. (2015). Explaining Speech Injustice: Individualistic vs. Structural Explanation. In D.C. Noelle, R. Dale, A.S. Warlaumont, J. Yoshimi, T. Matlock, C.D. Jennings, and P.P. Maglio (Eds.), *Proceedings of the 37 Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.
- [2] Kenny, A. (1963). *Action, Emotion and Will*. London: Routledge.
- [3] Marques, T. (2018). Retractions. *Synthese*, 195(8), 3335–3359. DOI: 10.1007/s11229-015-0852-8.
- [4] Marques, T. & García-Carpintero, M. (2020). Really Expressive Presuppositions and How to Block Them. *Grazer Philosophische Studien*, 97(1), 138–158. DOI: 10.1163/18756735-09701008.
- [5] Stanley, J. (2015). *How Propaganda Works*. Princeton: Princeton University Press.
- [6] Starr, W B. (2020). A Preference Semantics for Imperatives. *Semantics and Pragmatics*, 13(6). DOI: 10.3765/sp.13.6.

The way things go: Moral relativism and suspension of judgment

Eduardo Pérez-Navarro*
University of Granada

A popular accusation against moral relativism is that it goes too far in its vindication of tolerance. By setting no limits to what we should tolerate, the accusation goes, moral relativism condemns us to shrug our shoulders no matter what disgusting practice we are confronted with, for every practice is deemed permissible by some moral standard or another. The idea behind accusations like this can be summarized in the slogan, frequently attributed to relativism, that “anything goes”. The aim of this paper is to defend moral relativism from the accusation that it is an “anything goes” view; from the accusation that it forces us to *suspend our judgment* in cases in which we do not think we should even be allowed to. In the end, as we will see, relativism is not an “anything goes” view because it is not a view about *what goes*, but about *the way things go*—about what goes on when we say that something is morally right or wrong.

Here, I characterize moral relativism in terms similar to the ones contemporarily used to characterize the corresponding view about predicates of personal taste, which I borrow from MacFarlane (2014): moral relativism is the view that the truth of moral claims is relative to the *context of assessment*, that is, the context from which we wonder whether the claim is true or false. There is indeed a view, sometimes called “relativism”, that forces us to suspend our judgment about every practice, but it is not so much moral relativism as moral *contextualism*—the view that the truth of moral claims is relative to the *context of use*, or the context at which the claim is made. Once the context of assessment gets into the picture, however, we are deprived of the resources that would allow us to say that “anything goes”; the contextualist, for whom these resources are still available, is the one who is forced to say so. To see this, consider the following moral claims:

- (1) Female genital mutilation is morally wrong.
- (2) Female genital mutilation isn’t morally wrong.

Suppose that Alice utters (1) because she endorses a moral standard that condemns FGM, while Beth replies with (2) because her moral standard permits it. In cases like this,

* edperezna@gmail.com

contextualism says that (1) is true relative to Alice's standard and false relative to Beth's, while (2) is true relative to Beth's standard and false relative to Alice's. As outrageous as Alice might find FGM, hence, she should realize that its moral valence is relative to a moral standard, thus making her disagreement with Beth faultless (Kölbel 2004; Wright 2006) and tolerance the stance to take.

However, what Alice is interested in is not whether FGM is right or wrong according to this or that, but whether it is right or wrong, period. When deciding whether FGM is right or wrong, she should consider whether (1) and (2) are true or false *simpliciter*. For (1) to be true or false *simpliciter*, though, is for it to be so with respect to Alice's moral standard, while for (2) to be true or false *simpliciter* is for it to be so with respect to Beth's. (1) and (2) are thus both true *simpliciter*. All Alice could do in this situation is therefore to suspend her judgment about the matter.

Relativism allows Alice to be aware of the fact that (1) and (2) are both true relative to the standard of the person who utters them, thus making it easier for her to be tolerant, but this does not mean that, for her, FGM is neither right nor wrong. For this to be so, she would have to be able to put both her standard and Beth's into play when assessing the moral valence of FGM. But she only has the standard that she has, and she cannot "jump over it" to assess FGM from someone else's point of view. If she could, she would be able to see things as they stand from a perspective that makes Beth's standard just as valid as hers; but she simply cannot do this.

Apparently, though, the most salient alternative to "anything goes" views such as contextualism is not moral relativism. It is moral *objectivism*, according to which there is a fact of the matter about moral issues (see e.g. Hirvonen 2016). However, objectivism too ends up being an "anything goes" view unless the objectivist takes herself to be endowed with "God's point of view". To see this note, as a preliminary point, the equivalence between saying that the truth of moral claims is not relative to moral standards and saying that it is always relative to the same standard, which we may call m_o . This is the "objective standard", the one that would hold from God's point of view. Objectivism would then be the theory according to which (2), as uttered by Beth and assessed by Alice, is true if and only if FGM is morally right relative to m_o .

There are two options for the objectivist at this point. The first one is to take Alice's standard to be in fact *mo*. By doing so, we would obtain a kind of objectivism that we may call *absolutist*:

Absolutist objectivism: Alice should judge that Beth has said something true iff FGM is morally right relative to the objective moral standard.

If Alice's standard is m_o , the context that determines for Alice the truth of (2) is the same that determines for her the truth of (1), so she cannot suspend her judgment. However, this requires Alice to be assessing (2) from a context in which m_o holds, or, in other words, from God's point of view. I take this to be unacceptable for two different reasons. First, absolutist objectivism requires the postulation of a realm of "moral facts", together with the assumption that Alice, or any of us for that matter, is connected with them in a way that makes no room for error. If $m_A = m_o$, Alice should be infallible. This is too hard a pill to swallow. Second, absolutist objectivism fits badly with tolerance. It is easier to be intolerant if you think that you have the only correct standard, something that is more readily available for objectivists than for relativists (Prinz 2007: 208).

Of course, almost no one who holds that moral claims are true or false relative to an objective standard will take herself to be infallible or behave in an intolerant way. Objectivists who want to avoid these undesirable consequences have a second option available to them –*open-minded* objectivism:

Open-minded objectivism: Alice should judge that Beth has said something true iff she believes FGM to be morally right relative to the objective moral standard.

An open-minded objectivist would take moral claims to be always relative to m_o , but would accept that human nature is limited in such a way that m_o , even if the standard that we *should* use to assess our moral claims, is never the standard that we *do* use. If we were fully committed to the truth of a moral claim, we would be so from God's point of view. But, since such a point of view is unattainable, all our opinions on moral issues are provisional (see Rachels 1986: chapter 2).

Open-minded objectivism does not make us infallible, but avoids this at the cost of losing what little would make absolutist objectivism preferable to contextualism—it makes room for, and in fact forces suspension of judgment, since it describes matters in such a way that we can never be certain about the moral valence of any practice. It makes it possible for us to remain silent until we have attained God's point of view, something that, of course, never comes.

Objectivists thus face a dilemma. Either they countenance infallible agents and thus foster intolerance, or they are forced to accept suspension of judgment back. Relativism does not find itself in such a situation, for it can avoid and even forbid suspension of judgment without introducing God's point of view into the picture and throwing tolerance out of it. Relativism thus offers us effective weapons against the two-headed monster of contextualism and objectivism, whose objectivist head is itself two-horned.

References

- [1] Hirvonen, S. (2016). Doing without judge dependence. In C. Meier and J. van Wijngaert (eds.), *Subjective Meaning: Alternatives to Relativism* (pp. 47–68). Berlin: De Gruyter.
- [2] Kölbel, M. (2004). Faultless disagreement. *Proceedings of the Aristotelian Society*, 104(1), 53–73. MacFarlane, J. (2014). *Assessment Sensitivity: Relative Truth and Its Applications*. Oxford: Oxford University Press.
- [3] Prinz, J. J. (2007). *The Emotional Construction of Morals*. Oxford: Oxford University Press. Rachels, J. (1986). *The Elements of Moral Philosophy*. Philadelphia: Temple University Press.
- [4] Wright, C. (2006). Intuitionism, realism, relativism and rhubarb. In P. Greenough and M. P. Lynch (eds.), *Truth and Realism* (pp. 38–60). Oxford: Oxford University Press.

Metaphor identification via contextual coherence

Claudia Picazo Jaque*
Universidad de Granada

Inés Crespo**
NYU Paris

Andreas Heise***
Institut Jean Nicod (CNRS, EHESS, ENS-PSL)

In this paper, we propose to take up the problem of how to identify metaphors. We show that a principled analysis of the identification of metaphors is possible. In order to do this, we make use of the notion of Questions Under Discussion and offer a pragmatic approach.

Traditionally, the hallmark of metaphor was thought to consist in some sort of semantic anomaly, conflict, or deviance—a view summed up nicely in Nelson Goodman's (1976, 73) gloss of metaphor as a 'calculated category mistake'. However, category mistakes are neither necessary nor sufficient for metaphoricality. On the one hand, Timothy Binkley (1974) and Ted Cohen (1976) drew attention to the phenomenon of twice-true metaphors, a phenomenon exemplified by sentences (1) and (2):

- (1) Moscow is a cold city.
- (2) I have to get out, I can't breathe here.

(1) and (2) exhibit no semantic anomaly whatsoever. In other words, they lend themselves to a literal reading. Yet, as the case may be, a metaphorical interpretation is possible as well, hence the term 'twice-true' metaphors.

On the other hand, some category mistakes are best seen as plain category mistakes. We can think here of a child standing in front of the Law Faculty and uttering (3):

- (3) I can't see the university.

Although (3) involves some sort of semantic anomaly, it is not plausible to interpret the child as uttering a metaphor. Given that sentences may count as metaphors without being semantically anomalous, and some semantic anomalous sentences do not count

* claudia.picazo@gmail.com

** inescrespo@gmail.com

*** andreas.heise@gmx.ch

as metaphors, semantic anomaly fails to pick out reliably the phenomenon we are interested in.

Another option is to look for some kind of pragmatic conflict. In the Gricean view of metaphor, the metaphorical interpretation is derived as an implicature. This mechanism presupposes a pragmatic mismatch—the literal reading is inappropriate, and so the interpreter infers that the speaker means something else. However, twice-apt metaphors call into question this analysis. A twice-apt metaphor is a sentence that is to be interpreted both as a metaphor and as a literal claim (in a given context). Lynne Tirrell gives the following example:

Consider, for example, the case of the company president who looks out the boardroom window at a sky full of stratus clouds just before meeting with the company's union labor leader. The company president says to her assistant, [(6)] 'There's a storm brewing'. There is no incongruity between a literal interpretation of 'storm brewing' and this context [...]. But the figurative interpretation of 'storm brewing' is appropriate as well. (Tirrell, 1991, 355-6)

Instead of focusing on deviance or conflict, the key notion in our approach is that of contextual coherence. We hold that, in general, the output of an interpretation of a sentence-in-context must secure contextual coherence. This idea can be applied to metaphors. For this, we introduce the notion of a minimal paraphrase and a specific notion of context.

The minimal paraphrase of a metaphor is not a paraphrase of the literal meaning of the sentence, but something that hints at the expressive power of the metaphor. It captures the gist of the metaphor without exhausting it. We can think of it as the kind of paraphrase that we would offer to someone who has not grasped the metaphor. Non-Cognitivists about metaphorical meaning like Donald Davidson (1978) might deny that metaphor involves anything like a minimal paraphrase that goes beyond literal meaning. However, there are two reasons for thinking that there is such a minimal paraphrase in the case of metaphor, and that the paraphrase fits the format of propositions. First, we deem it reasonable to think that understanding a metaphor involves being able to explain, roughly, what the metaphor says, using other words. Second, the minimal paraphrase is needed to deal with some linguistic phenomena as implicatures derived from metaphors.

We offer the following definition:

[Metaphor] A token of a sentence is reasonably seen as a metaphor with minimal paraphrase *mp* if and only if *mp* is coherent with the context.

This allows us to apply to the study of metaphor a tool that has proved fruitful on accounting for other phenomena in natural language: Questions Under Discussion (QUD) (see e.g. van Kuppevelt 1995 and Roberts 1996). QUD allow us to capture the idea that

the interpretation of a sentence usually depends on the topic of the conversation, they afford a nice way to capture the notion of ‘topic of the conversation’. In our view, QUD are part of the context, conceived as (a part of the) interpreter’s epistemic state. As we see it, the context includes previous discourse, information about surroundings (such as perceptual information) and background information (for example, about literary genres). QUD can be explicitly stated (thus being part of the previous discourse), introduced via an assertion or otherwise inferred from the three types of information mentioned.

Contextual coherence involves two ingredients. First, in order for an interpretation to be coherent it must be an answer to a relevant QUD. With this we are able to account for some cases of metaphor identification. Imagine the following conversation:

A: Have you met some nice people in Moscow?

B: Moscow is a cold city.

A metaphorical reading of (1) is an (indirect) answer to the QUD introduced by A. Thus, it secures coherence. (By contrast, a literal interpretation would be off-topic).

Second, we add a requirement of external consistency. External consistency is understood as consistency with contextual information. This allows us to deal with more examples:

C: How did you like Moscow?

D: Moscow is a cold city.

Both the literal and the metaphorical readings are compatible with the QUD. However, suppose that we have additional information: C knows that D comes from the Russian Taïga, where temperatures are far lower than in Moscow. In this case, the literal interpretation is ruled out, the metaphorical interpretation is preferred.

How do the two conditions on coherence sum up? Condition 1 is necessary but not sufficient; condition 2 is also necessary but not sufficient, and it is secondary to condition 1. This means that an interpretation might be incoherent either when it does not address at all an issue in the QUD, or when it does but when it is not externally consistent. The two conditions are jointly sufficient.

Our view allows for more than one interpretation being coherent. For instance, there might be a plurality of relevant QUD at a certain point of a discourse or conversation. This explains Tirrell’s example. In the context described, it is plausible to suppose that the context involves two QUD, one about the meeting and another about the weather. The metaphorical interpretation addresses the first, whereas the literal interpretation addresses the second. Also, there can be degrees of coherence. A reading might be reinforced or weakened by contextual information, thus being more or less coherent.

References

- [1] Binkley, Timothy. 1974. "On the Truth and Probity of Metaphor." *Journal of Aesthetics and Art Criticism* 33 (2): 171-80.
- [2] Cohen, Ted. 1976. "Notes on Metaphor." *The Journal of Aesthetics and Art Criticism* 34 (3): 249-59. <https://doi.org/10.2307/430007>.
- [3] Davidson, Donald (1978), "What Metaphors Mean," *Critical Inquiry* 5: 31-47
- [4] Goodman, Nelson. 1976. *Languages of Art: An Approach to a Theory of Symbols*. Indianapolis IN: Hackett.
- [5] Roberts, C. (2012). Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics & Pragmatics*, 5, 1-69.
- [6] Tirrell, L. (1991). Reductive and Nonreductive Simile Theories of Metaphor. *Journal of Philosophy*, 88 (7), 337-358.
- [7] van Kuppevelt, J. (1995). Discourse structure, topicality and questioning. *Journal of Linguistics*, 31, 109-147.

Tracing Thick Concepts Through Corpora

Kevin Reuter*
University of Zurich

Lucien Baumgartner**
University of Zurich

Pascale Willemsen***
University of Zurich

Thick concepts, such as *rude* and *courageous*, describe and evaluate at the same time. Despite the ubiquity and importance of thick concepts, we lack a reliable means to tell thick concepts apart from other concepts, including descriptive concepts (e.g., *blue*), purely evaluative (that is, thin) concepts (e.g., *terrible*), and concepts, such as *homeless*, that are value-associated but do not evaluate. No tool for measuring the polarity and evaluative intensity of a thick concept has so far been developed. Instead, scholars rely almost exclusively on their own intuitions in order to identify thick concepts, usually sticking with a list of paradigmatic examples that most people agree with (Roberts, 2013; Väyrynen, 2013).

Background

Given the limitations of intuitive classification of concepts, we propose to approach the identification and measurement of thick concepts using tools from corpus linguistics. We present the results of two corpus-linguistic studies for a wide range of thick, thin, value-associated, and descriptive adjectives. In Study 1 we show that evaluative adjectives behave differently (compared to descriptive adjectives) when in combination with other adjectives through the connective ‘and’. Study 2 reveals that the modifiers ‘truly’ and ‘really’ highlight the evaluative dimension of thick and thin concepts, allowing for them to be reliably classified.

* Kevin.Reuter@philo.unibe.ch

** lucien.baumgartner@philos.uzh.ch

*** pascale.willemsen@uzh.ch

Study 1: And Connective

We selected 45 adjectives, to be investigated in both studies. Given the function of the term ‘and’, we can hypothesize that thick positive terms, descriptive terms, and thick negative terms can be differentiated from each other merely by analyzing the words with which they are combined using the connective ‘and’. We made the following predictions (see Table 1 below) for negative thick adjectives (nta), positive thick adjectives (pta), and descriptive adjectives (da):

| Combination | Valence of x | Examples |
|-------------|-------------------|---|
| “nta and x” | negative, neutral | ‘rude and immature’, ‘rude and demanding’ |
| “pta and x” | positive, neutral | ‘friendly and playful’, ‘friendly and conversational’ |
| “da and x” | any | ‘circular and blue’, ‘loud and silly’ |

Table 1: Predicted Valence of Conjoined Adjectives.

The first corpus study is based on 223,403 Reddit comments we collected via the pushshift.io API. We fixed a quota of at least 2000 conjunctions for each target adjective. To determine the valence of an adjective that is combined with a target adjective, we used sentiWords, a freely available dictionary containing prior polarity scores for many words. The numbers range from ‘-1’ meaning ‘absolutely negative’ to ‘+1’ meaning ‘absolutely positive’. The observed and estimated means are depicted in Figure 1. We computed pairwise comparisons between the means of the polarity classes using estimated marginal means (EMM). These comparisons demonstrate significant differences among all three groups: negative - neutral: $t(223400) = -102.510$, $p < 0.001$; negative - positive: $t(223400) = -282.736$, $p < 0.001$; neutral - positive: $t(223400) = -107.779$, $p < 0.001$.

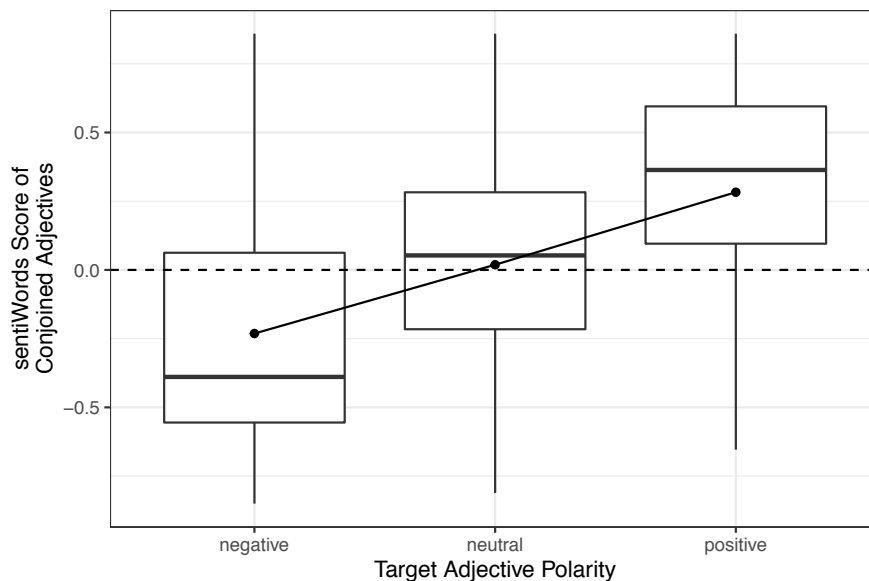


Figure 1: Boxplots of observed sentiment values.
The dots represent the estimated means per polarity group.

The three main predictions of Study 1 were supported by the data we collected and analyzed. Through the connective ‘and’, positive thick adjectives combine more frequently with positively-laden adjectives, whereas negative thick adjectives do so with negatively-laden adjectives. Additionally, adjectives that are combined with descriptive adjectives have a significantly lower absolute sentiment value.

Study 2: Truly and Really Modifier

Whereas the true-modifier can be used to stress the normative dimension of dual character concepts, as in ‘true mechanic’, the modifiers ‘truly’ and ‘really’ seem to intensify the evaluative aspect of thick adjectives, as in ‘truly courageous’ and ‘really selfish’. Consequently, we predicted that ‘truly x’ is more common for thick and thin adjectives compared to descriptive and value-associated adjectives. In other words, the ‘truly’ modifier allows us to distinguish those classes of concepts that have an evaluative dimension (thick and thin concepts) from those that do not (descriptive and value-associated concepts).

The 45 concepts used in Study 1 were examined again in Study 2. We decided to use the Corpus of Contemporary American English (COCA) for this task. For all 45 concepts, we recorded the amount of hits for ‘truly x’, ‘really x’, and ‘x’ on COCA. We then calculated the ratios (e.g., ‘truly x’ divided by ‘x’) and normalized them for all 45 concepts. The values we received for all 45 adjectives show that almost all value-associated and descriptive terms had lower values than thick and thin terms. Only for ‘rich’ and ‘loud’ did the modifier approach yield results that put them above some thick terms. We

also performed a hierarchical cluster analysis using squared distance (Ward’s method) to identify the inherent structure of the data. The results are displayed in the form of a tree diagram (Figure 2). The cluster analysis yielded three main clusters. In the ‘descriptive’ cluster (middle) only descriptive as well as value-associated concepts were included (not a single thick or thin term). The two ‘evaluative’ clusters (left and right) included only two terms that were originally classified as descriptive or value-associated (‘loud’ and ‘rich’; colored in light-grey in Figure 2 below).

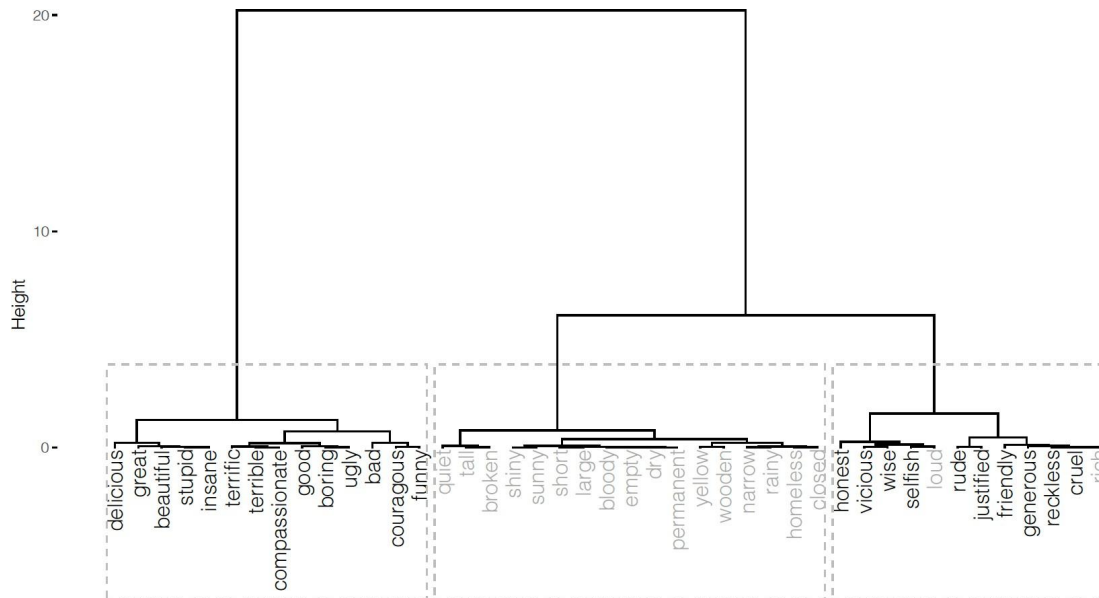


Figure 2: Tree diagram displaying the clusters using hierarchical cluster analyses.

We consider the results that demonstrate a separation of evaluative adjectives from value-associated adjectives particularly encouraging. In sum: We have derived stable and plausible results using data that reveals how the ‘truly’ and ‘really’ intensifiers work. And trying out different operationalizations, like the ‘and’ connective method, seems to have provided additional support for our claims.

General Discussion

Study 1 and Study 2 were designed to measure the evaluative dimension of concepts. In this paper we have sketched what an operationalization and measurement of evaluative intensity could look like. Let us focus on the operationalization and measurement results of Study 2, as we believe the outcome of this study is probably the most promising in terms of measuring evaluative intensity.

In Study 2, we proposed that a good indicator of or proxy for the evaluative intensity of a term is the extent to which the intensifiers ‘truly’ and ‘really’ can be reasonably applied to that term. We then operationalized that proxy through the ratio between the frequencies with which a term is intensified by ‘truly’ and ‘really’ and the overall frequency of that term, leading us to the variable *eval*. The results revealed a rather differentiated picture, according to which ‘truly’ and ‘really’ are most reasonably applied to thin concepts and thick concepts, and not very reasonably applied to value-associated and descriptive concepts. Figure 3 shows the probability that an adjective is evaluative based on its *eval* value (logit model). This graph shows that adjectives that have an *eval* value that is greater than 0.75 are quite likely to be thick or thin adjectives. Below an *eval* value of 0.3, adjectives are far more likely to be descriptive or value-associated. Unfortunately, our design does not allow us to say whether adjectives that receive low *eval* numbers are evaluative to a very low degree (or in very few contexts), or whether there is a threshold that distinguishes value-associated from evaluative terms.

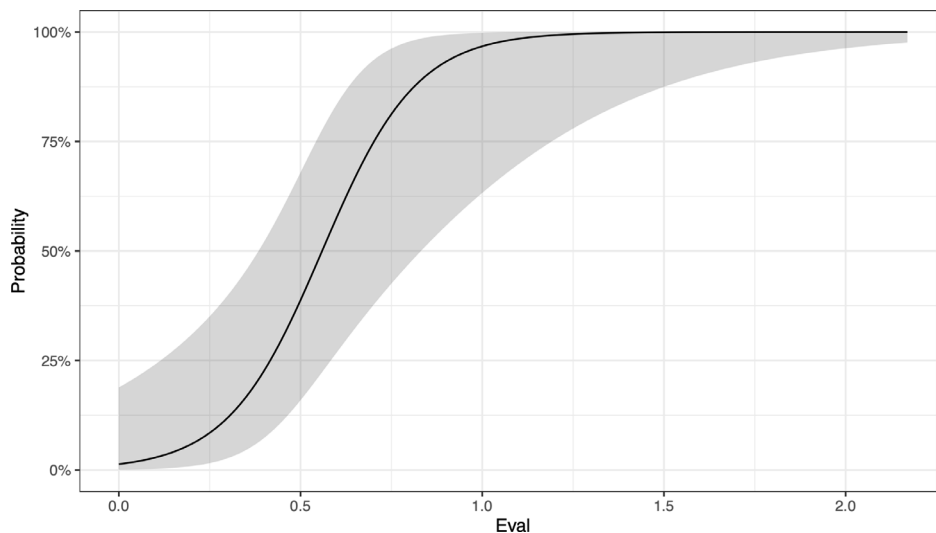


Figure 3: Predicted probability for being an evaluative concept, with confidence intervals.

Now, the crucial question is: Are we justified in saying that *eval* tells us the evaluative intensity of a term? Here are four reasons to answer this question in the affirmative, if tentatively:

1. We have motivated the operationalization of evaluative intensity through *eval* independently of the results we collected.
2. A cluster analysis demonstrated that *eval* allows us to match most pre-theoretic intuitions on the level of classes of concepts.
3. Almost all value-associated concepts received very low *eval* values.
4. For thin concepts, *eval* values matched the semantic meanings of the terms: ‘terrific’ and ‘awful’ are more evaluative than ‘good’ and ‘bad’, and correspondingly received higher *eval* values.

References

- [1] Roberts, D. (2013). Thick Concepts. *Philosophy Compass*, 8(8), pp. 677-688.
- [2] Väyrynen, P. (2013). *The Lewd, the Rude and the Nasty*. Oxford: Oxford University Press.

Logical conventionalism, truthmaking and expressivism

Pablo Rychter*
University of Valencia, Spain

1 Introduction

This short paper aims to contribute to clarifying and assessing *logical conventionalism*, focusing on the dialectics between Ted Sider (2011) and Amie Thomasson (2020). A critic of logical conventionalism, Sider argues that the idea of truth by convention is helpless. On the other side, Thomasson argues that Sider’s criticisms are misguided since they rely on saddling logical conventionalism with the implausible view that truth by convention is truth *about* conventions. I argue that, in spite of their disagreement, both Sider and Thomasson are mostly right. My assessment relies on two distinctions that I draw: one between the *truthmaker* of a sentence and what the sentence is a *about*, and the other between two core ideas that underly logical conventionalism: that logical truths are true by convention, and that logical “truths” do not aim to describe the world. The first distinction is required, I argue, to see the initial plausibility of Sider’s arguments. It is only if we fail to make that distinction that Sider’s arguments may seem misplaced. The second distinction is required to see that why, as Thomasson suggests, the criticism of truth by convention do not reach a central insight of logical conventionalism: that logical “truths” are not really true because they merely *express* something by saying nothing.

2 Two core ideas of logical conventionalism

It seems initially plausible that the following sentence is true:

(1) Either we are or are not in Salamanca.

And it may be natural to think the reason why (1) is true is that we happen to be in Salamanca. But the logical conventionalist thinks otherwise: the truth of (1) does not *depend* on the fact that we are in Salamanca, and this fact is not the *reason* why (1) is true. Instead, (1) is true *in virtue of* the conventions that we language users adopt and that define the meaning of ‘either..or...’. This dependence of logical truth on convention is a

* pablo.rychter@uv.es

first core idea logical conventionalism, an idea that we could also express, as Ted Sider does, in terms of the notion of *truthmaking*:

The core of the view is that an analytic truth, for instance the truth that all horses are horses, is true purely by virtue of linguistic conventions. By adopting certain rules governing the use of logical words like ‘all’, **language users somehow make ‘all horses are horses’ true.**” (Sider 2011: 98, emphasis added).

A **second core idea** is that logical truths, because they only ‘record’ or ‘express’ linguistic conventions, are *devoid of factual content*. As Ayer puts it, the propositions of logic ‘do not make any assertion about the empirical world, but simply record our determination to use symbols in a certain fashion’.

3 Sider’s criticism of logical conventionalism

Sider’s critical discussion of logical conventionalism takes Quine’s “Truth by convention” as its starting point. Sider distinguishes two arguments in Quine: the first, which we are not going to discuss, is that logic conventionalism incurs in some sort of regress: we need logic already to set up our conventions about language use, and so logic cannot rely on conventions. The second is that the idea of *truth by convention* cannot be satisfactorily explained. Sider considers different attempts at explaining the idea and finds all of them wanting. For instance, it is unsatisfactory to just say (1) does not require the *world’s cooperation* to be true, or that its truth does not impose any *requirement on the world*. This is because (1) *does* impose the relatively easy-to-satisfy requirement that we either are, or are not, in Salamanca. Sider concludes that there is no clear sense in which logic could be true by convention. His discussion vindicates Russell’s view that “logic is concerned with the real world just as truly as zoology, though with its more abstract and general feature”.

4 Thomasson’s defense of logical conventionalism

According to Thomasson, Sider’s discussion of logical conventionalism misconstrues its target. Its main mistake is to attribute conventions a truth-making role: “Taking conventionalism to be the view that conventions serve as the truthmakers for modal statements (...) is historically inaccurate and philosophically catastrophic” (2020: 27). Thomasson thinks that this misconception that leads to what we may call *the standard objection*: that logical conventionalism makes logical truths *contingent*, as much as our adoption of certain conventions is, and that it is therefore unable to explain the felt necessity of logical truths. But, Thomasson argues, this is a mistake. In her words:

To take all those who are labeled as “conventionalists” to be defending the view that the necessary truths of logic and mathematics are descriptive claims made true by our adopting certain linguistic conventions ignores the really interesting and promising aspect of their proposal, and of the Wittgensteinian view that inspired it—namely, that the propositions of mathematics and logic should not be thought of as descriptive claims in need of truthmakers at all. (2020: 27)

5 Discussion and conclusions

In order to assess the arguments of Thomasson and Sider, it will be helpful to focus on the following argument.

- P1. According to logical conventionalism, logical truths are *made true* by conventions
- P2. If e makes p true, p is *about* e (or p *describes* e).
- C. According to logical conventionalism, logical truths are about (or describe) conventions.

Although this argument is not explicitly stated by Thomasson, I think it aptly represents the reasoning she attributes to Sider. Notice that the conclusion C leads to the *standard objection* described above: if logical truths describe conventions, then they are as contingent as conventions are, which makes logical conventionalism implausible. However, I think that Sider neither subscribes C, nor the *standard objection*. (He considers a “nonstarter” the suggestion that analytic truths make statements about linguistic conventions (2011: 101). So I think there is a little bit of a misattribution here on behalf of Thomasson). In any case, I think it is clear that logical truths are not *about* conventions, and that Thomasson is right that C is a misinterpretation of logical conventionalism. However, Sider does accept P1, which according to Thomasson, as we saw above, is the culprit of all misinterpretations. Should we accept P1?

On the one hand, I think we could accept P1, since it is a reasonable understanding of the **first core idea of logical conventionalism** that we distinguished above, the idea of *dependence* of logical truth on convention. I disagree with Thomasson that P1 is clearly a misconstruction of logical conventionalism. The idea attributed by C (that logical truths are about conventions) is a nonstarter, but not so the idea attributed by P1 (that logical truths are made true by convention). But if we accept P1 and reject C, we should of course reject P2, an assumption that Thomasson makes all along and takes her back and forth from P1 to C. In spite of Thomasson reliance’s on it, I think P2 is false. A truth need not describe or be about its truthmakers: I make true the proposition that there are human beings in Salamanca, but the proposition is arguably not *about* me, and does not describe me. Once we reject P2, we can then use the idea of truthmaking to understand logical

conventionalism, as Sider does, without incurring the misconception expressed by C and the *standard objection* that relies on it.

On the other hand, perhaps we should reject P1, as Thomasson urges us to do. This is what we should do if we wanted to focus on the **second core idea** of logical conventionalism distinguished above, and elaborate the idea that the propositions of logic are *not descriptive* in terms of the idea that they are *not true*. On this suggestion, it is wrong to see conventions as truthmakers of the propositions of logic, because those propositions are simply not true. They are also not false, since they are best understood as expressing something, without really asserting anything. This is of course an alternative, expressivist development of logical conventionalism, initially incompatible with idea that the propositions of logic are *true* by convention. This alternative development is unaffected by the objections against truth by convention, but faces other difficulties more directly, as the Frege-Geach problem, or the simple complaint that (1) *does* seem true.

Let us summarize. First, if we distinguish between being a truthmaker and being object of description, Sider's criticisms are on target against the idea of truth by convention and do not saddle logical conventionalists with the idea that logical truths are *about* conventions. (Thomasson misconstrues Sider's view when she argues that he misconstrues the conventionalist's views.). Second, Thomasson is right that there is an alternative construal of logical conventionalism, based on the second core idea, which is unaffected by Sider's criticism.

References

- [1] Sider, Theodore (2011). *Writing the Book of the World*. Oxford University Press.
- [2] Thomasson, Amie L. (2020). *Norms and Necessity*. Oxford University Press..

Are empty names all the same?

Mirco Sambrotta*
Slovak Academy of Sciences

Some (most) attempts to introduce proper names succeed, but it can happen that some names that are introduced with the intention to refer to something real never become proper names because there is no object to refer to. A well-known example of dubbing/baptism attempt failure is the case of ‘Vulcan’, which the French astronomer Le Verrier introduced in order to name a hypothetical intra-Mercurial planet, but which then turned out to refer to nothing. Nevertheless, I will suggest, even though ‘Vulcan’ did not succeed to refer to any object, that does not imply that it did not name anything at all. Indeed, ‘Vulcan’ has become established as a common currency word independently of the fact that it did not turn out to be a proper name.

What I will try to put forward is then that, given certain required circumstances and conditions are fulfilled, a name that does not refer to a (modally robust) individual, although cannot be a proper name, can still denote a concept. What ‘Vulcan’ stands for is thus the very description originally (unsuccessfully) used to fix the reference of the name, i.e., ‘The intra-Mercurial planet causing the perturbations in the orbit of Mercury’ or something like that. In this way, ‘Vulcan is the intra-Mercurial planet causing the perturbations in the orbit of Mercury’ can be understood as covertly metalinguistic, a definition, an analytic statement, or a grammatical proposition, but not as a factual proposition. It follows that ‘Vulcan is inhospitable, with no trace of life’ should be understood as expressing the (false) particularized proposition ‘The intra-Mercurial planet causing the perturbations in the orbit of Mercury is inhospitable, with no trace of life’.

Nevertheless, Le Verrier really believed his dubbing/baptism attempt succeeded and ‘Vulcan’ stood for a proper name, therefore it seems plausible to attribute to him the false belief that ‘Vulcan is the intra-Mercurial planet causing the perturbations in the orbit of Mercury’ expresses a factual proposition instead of a grammatical one, as well as the false belief that ‘Vulcan is inhospitable, with no trace of life’ expresses a singular proposition instead of a (false) particularized one. According to this view, ‘Vulcan’ thus stands for an (empty) description outside the intensional (or opaque) context of the propositional attitude, but inside the intensional context, it is falsely believed to stand for a proper name—albeit it is not a proper name at all.

Now, one might be tempted to think that this very idea that empty names are synonymous with certain descriptions is even more fitting in the case of fictional names, for in

* mirco.sambrotta@gmail.com

the case of a fictional character the description that the name abbreviates would simply be the conjunction of all the properties that in the story the character is said to possess. In this way, as role realists suggest, fictional characters turn out to be roles or offices constituted by sets of (necessary) requisites (Wolterstorff 1980; Currie 1988, 1990; Lamarque and Olsen 1994; Stokke 2020; Glavaničová 2018, 2021). However, what role realists seem to overlook is that by means of stories authors pretend to name fictional characters and state propositions about them. That is to say, they imagine (and let us imagine) such names to behave as proper names—and exactly this attitude is what is made explicit by modifier or story operators (or in-the-fiction operators), such as ‘in the fiction...’, ‘according to the story...’, etc., understood as higher-order operators.

As well as the previous example, also work of fictions (about fictional characters) can be understood as manifolds of elucidatory pseudo-sentences that purport to introduce and define concepts, but which are imagined or pretended as sentences expressing factual propositions about certain individuals. Even though ‘Sherlock Holmes is a detective’ turns out to be a covertly metalinguistic claim by means of which Conan Doyle introduced the concept Sherlock Homes, defining it as having among its necessary requisites that of being a detective, at the same time, inside the pretense of the story, it is imagined as expressing a factual singular proposition.

To make things even more complicated, we could also imagine someone writing a fictional story about a non-existent object, whose name is already a common currency word, for instance, Vulcan (there are many indeed). Would that be a fictional story about a possible object or a fictional story about a fictional object? Or even, we could imagine a whole story about a certain non-existent object, but which the author really believed to exist, so that her intention was that of writing a historical report about it. Would that count as a work of fiction? In this talk, I will contend that in order to address such issues, we should focus on the attitudes involved, that is, the author’s intention to write a work of fiction or not always plays a central role in accounting for and distinguishing fictional stories from false reports. Her intention will mark the distinction, on the condition, though, that it is recognizable by the audience in the relevant context. By this I mean that an intentional act should not be understood as something that happens within the mind of the author or anyone else, but rather as a socially and inferentially regulated public act.

In a nutshell, what I will try to show is that empty names are not proper names at all, they are instead better understood as empty disguised individual descriptions (with concepts as intensions, but without extensions) that in intensional context are taken as standing for proper names. As such, an empty name is an expression that has no denotation in the actual world but could have it in some possible worlds, that is, it picks out no actual individual, but it could pick out some possible individuals (in the possible worlds in which there is whatever individual that meets the description that the name abbreviates) albeit it would not pick out any specific individual in any possible world. It could be taken as it does, as it picks out a specific individual, just when it is used in an intensional

context—albeit there is not and it is impossible there could be such an individual. If we like, we could call the former ‘possible objects’ and the latter ‘impossible objects’, but we should keep in mind that in doing so we are not committing ontologically with any object, but we are just introducing a deflationary way of speaking. The attitude that embeds the sentence in which they occur will then determine which kind of impossible object they are (if hypothetical, fictional, mythological, and so on).

Accordingly, in contrast to realism, one of the main goals of this talk is to advance a naturalistic account of empty names. As such, it aims to avoid any ontological commitments with respect to entities that cannot be conformed to the explanations provided by science. Of course, one could object that although it is not committed with any spooky individuals (such as hypothetical objects or fictional objects), it is anyway committed with abstract entities. In fact, according to this view, empty names are not proper names that refer to individuals, but rather disguised descriptions that pick out bundles of properties. Nevertheless, unlike role realists, talk about abstract entities here is not understood substantively but rather, following Sellars (1963), as talk about linguistic types. Or better, what we are doing in talking about abstract objects is codifying, reifying, nominalizing discursive norms. This may look like a hard pill to swallow, especially to accommodate cases where we seem to be making true assertions about non-existent objects and which do not allow the move of adding intensional operators —e.g., ‘Sherlock Holmes is a fictional character’ or ‘Vulcan doesn’t exist’. But, in the end, the pill turns out to be not so bitter if here ‘to be a fictional character’ and ‘to exist’ are nothing but second-order properties. As a result, they can be re-formulated meta-linguistically —e.g., “‘Sherlock Holmes’ is a proper name in a fictional story’ or better “‘Sherlock Holmes’ is a predicate that behaves as a proper name in a fictional story’ and ‘Vulcan has no extension’ or even ‘The intension corresponding to Vulcan does not single out any individual’.

References

- [1] Currie, G. (1988). ‘Fictional Names’. *Australasian Journal of Philosophy*. 66, pp. 471-488.
- [2] — (1990). *The Nature of Fiction*. Cambridge: Cambridge University Press.
- [3] Glavaničová, D. (2018). ‘Fictional Names and Semantics: Towards a Hybrid View’, in Stalmaszczyk, P.(ed.), *Objects of Inquiry in Philosophy of Language and Literature: Studies in Philosophy of Language and Linguistics*. Berlin: Peter Lang, pp. 59–73
- [4] — (2021). ‘Rethinking Role Realism’, *British Journal of Aesthetics* 61 (1):59-74
- [5] Lamarque, P. and Olsen, S. H. (1994). *Truth, Fiction, and Literature: A Philosophical Perspective*. Oxford: Clarendon.

- [6] Sellars, W. 'Abstract Entities'. *Review of Metaphysics* 16 (1963): 627–71.
- [7] Stokke, A. (2020). 'Fictional Names and Individual Concepts'. *Synthese*. Available at: <<https://doi.org/10.1007/s11229-020-02550-1>> (Accessed: 28 September 2020).
- [8] Wolterstorff, N. (1980). *Works and Worlds of Art*. Oxford: Clarendon.

Slurring Terms in Narrative

Isidora Stojanovic*

Institut Jean-Nicod, CNRS + Universitat Pompeu Fabra

Abstract

Philosophers' interest in slurring terms, in particular racial epithets such as the N-word for Black people, has given rise to a large literature over the past two decades. Somewhat surprisingly, occurrences of slurring terms in fictional narratives have been largely ignored. Our aim is to fill this lacuna. We are especially interested in slurs as they occur in so-called *free indirect discourse* (FID), a narrative style of reporting speech and thought that combines third-personal narration with direct, first-personal discourse. We shall provide examples from the literature in which slurs occur in FID. We shall then use these data to assess, on the one hand, the two main approaches to FID (namely, the two-context approach and the mixed-quotation approach), and, on the other, several recent accounts of slurs.

Warning: This chapter cites examples that contain offensive vocabulary. We apologize for any offense caused.

1 Introduction

There are two well-known ways of reporting speech and thought: direct discourse (henceforth DD) and indirect discourse (henceforth ID), illustrated in (1) and (2):

- (1) Iñaki said/thought: “I will call her tomorrow”.
- (2) Iñaki said/thought that he would call her on the next day.

In narrative, however, there is yet a third, often preferred way: *free indirect discourse* (henceforth FID), illustrated in (3):

- (3) Iñaki's mom was expecting his call, but it was late. He would call her tomorrow.

FID has features in common with both DD and ID. As in ID, certain referential devices such as pronouns and tenses require readjustment. Both in (2) and (3), the first person pronoun “I” from Iñaki's original utterance or thought is converted into the third person pronoun “he”, and the future tense “will call” takes on the past form “would call”. As in DD, on the other hand, other expressions, such as “tomorrow”, are preserved. When we read (3), we understand that Iñaki said/thought that he would call his mom on the day after his speech or thought took place, rather than the day after the narration in (3) takes

* isidora.stojanovic.nicod@gmail.com

place. In FID, indexicals such as “tomorrow” pick out their referent from the perspective of the person whose speech or thought is being reported, rather than the reporter’s.

While a large amount of the literature on FID has focused on the behavior of context-sensitive expressions such as “tomorrow”, another remarkable feature of FID is that it features a range of expressions and constructions that are preserved, if at all, in DD, but not in ID: expressive elements such as “damn” or “asshole”, various discourse particles, question- and exclamation marks, nonstandard pronunciation and dialect (indicated by different spelling), as well as disfluencies and pauses (indicated by typography). This paper focuses on a subclass of expressive terms, namely, slurs, such as the N-word. Our goal is to examine how slurs in FID impact the scope and the plausibility of the existing accounts both of FID and of slurring terms.

2 Slurs in Free Indirect Discourse

Slurs have a strong expressive component. In using them, the speakers express their hatred, contempt, or other negative attitudes toward not only the persons for whom the slur is used on a given occasion, but toward the entire group to which they belong. Since slurs are expressive, and FID likes to preserve expressive elements of all sorts, one should, in principle, expect to see occurrences of slurs in FID. *In principle*, because the offensive and derogatory character of slurs could be so strong as to overthrow a writer’s desire for the use of such expressive vocabulary in FID. We show that the initial expectation is borne out: we do find slurs in FID. For reasons of space, we will simply lay out three examples, without further comments.

Maycomb was interested by the news of Tom’s death for perhaps two days; two days was enough for the information to spread through the county. “Did you hear about?... No? Well, they say he was runnin’ fit to beat lightnin’...” To Maycomb, Tom’s death was typical. Typical of a *nigger* to cut and run. Typical of a *nigger*’s mentality to have no plan, no thought for the future, just run blind first chance he saw. (...) Just shows you, that Robinson boy was legally married, they say he kept himself clean, went to church and all that, but when it comes down to the line the veneer’s mighty thin. *Nigger* always comes out in ‘em. (Lee 1960, p. 244)

Now the boys were approaching the dead man, warily. They were joking and snickering out of fear. An unemployed psychologist, Jack understood this. (...) Each time they backed away, laughing, but then came a bit closer on their next approach. Vultures, Jack thought. Damn *nigger* vultures. But he chided himself for this thought. (Johnson 1996, p. 1)

[W]hen the next day his boss apologized for losing control, Art said he understood. And when his boss said that Art shouldn’t take what he said personally, in fact he knew Art was not a *Jap*, but a *chink*, plus he had called someone else a lazy *wop* just that morning, it was just his style, Art said again that he understood. (Jen 1994, p. 17)

3 Discussion

In Section 1, we presented the characteristic features of FID: expressives as well as shiftable indexicals are typically anchored to the protagonist's rather than the narrator's perspective, while pronouns and tenses behave the other way round. In Section 2, we presented empirical evidence, drawing on examples from fictional narratives, that shows that slurs in FID tend to be anchored to the protagonist's rather than the narrator's perspective. The goal of this last section is to see how these data fit into our best theories of FID, on the one hand, and of slurs, on the other.

3.1 Two approaches to FID

This section presents two competing approaches to FID. The *two-context* approach relies on the idea that the interpretation of FID requires relativization to two contexts: the narrator's vs. the protagonist's context. This idea is elaborated in different ways in Banfield (1982), Doron (1991), Schlenker (2004), Sharvit (2008), Eckardt (2012, 2015) and Abrusán (2020). The accounts differ on how they incorporate the two-context idea into semantics, but, simplifying matters considerably, the main point on which they all agree is that pronouns and tenses are interpreted with respect to the narrator's context, while shiftable indexical as well as various expressive elements, with respect to the protagonist's.

While the two-context approach stresses the analogies between FID and ID, the *mixed quotation* approach, prominently defended in Maier (2015), stresses the analogies between FID and DD. The main idea is that although quotation marks may not be visible, the text or discourse in FID *is* quoted, while certain elements, such as pronouns and tenses, may be *un*-quoted.

3.2. Slurs in FID speak in favor of the mixed quotation approach

In this final section, we point out two features that are peculiar to slurs, and then argue that because of them, the mixed-quotation approach accounts better for slurs in FID than the two- context approach does. The two features go by the terms *hyper-projectivity* and *complicity*. Hyper-projectivity stands for the observation that the derogatory content of a slur projects from *any* kind of embedding, even under speech- and attitude-reporting verb such as “says/thinks”, which are known to behave as “plugs”; that is, which can block a presupposition from projecting. *Complicity* stands for the observation that the derogatory content of a slur does not simply concern the attitudes of the speaker, but also of the hearer, and even by-standers. Recall that the two-context approach stresses the similarities between FID and ID; Sharvit (2004) even posits covert operators that shift the interpretation of certain expressions to the protagonist's context. True enough, those operators are different from the standard speech and attitude reporting verbs, given that

the interpretation of indexicals such as “tomorrow” is sensitive to the former but not to the latter. But the two-context approach tells us that when an expression whose interpretation normally does not shift under speech and attitude verbs gets anchored to the protagonist’s context in FID, then this is because the expression at stake is indexical, or speaker-oriented. The two- context approach can, then, handle the hyper-projectivity of slurs in a straightforward way if it can show that slurs, too, are indexical.

The latter, however, is a problematic assumption. There is a remarkable difference between utterances that explicitly express the speaker’s negative attitudes toward a group and those that contain slurs. This feature of slurs is known as “complicity”. As Camp puts it, “slurs make recalcitrant hearers feel complicit in the speaker’s way of thinking” (2013: 333). The difficulty for the two-context approach to account for the data is that there is no account of slurs on which it can immediately build so as to yield the right empirical predictions. The mixed-quotation approach therefore appears to have advantage over the two-context approach. The reason is that if there is any mechanism that can “quarantine” the derogatory content of a slur, then it is quotation. This allows the mixed-quotation approach to explain why the derogatory content of those occurrences of slurs in FID does not need to engage the narrator, let alone the writer, but can instead reflect the protagonist’s perspective.

Selected references

- [1] Abrusán, Márta (2020). “The Spectrum of Perspective Shift: Free Indirect Discourse vs. Protagonist Projection”, *Linguistics and Philosophy*. <https://doi.org/10.1007/s10988-020-09300-z>
- [2] Banfield, Ann (1982). *Unspeakable Sentences: Narration and Representation in the Language of Fiction*. London: Routledge.
- [3] Camp, Elisabeth (2013). “Slurring perspectives”, *Analytic Philosophy* 54 (3): 330-349.
- [4] Doron, Edit (1991). “Point of view as a factor of content”, in Moore, S., Wyner, A.Z. (eds.): *Proceedings of SALT 1*. NY, Cornell University: 51- 64.
- [5] Eckardt, Regine (2015). *The Semantics of Free Indirect Discourse: How Texts Allow Us to Mind-read and Eavesdrop*. Brill, Leiden/Boston.
- [6] Maier, Emar (2015). “Quotation and unquotation in free indirect discourse”, *Mind & Language* 30: 345-373.
- [7] Schlenker, Philippe (2004). “Context of thought and context of utterance (a note on free indirect discourse and the historical present)”, *Mind and Language* 19(3). 279-304.

- [8] Sharvit, Yael (2008). “The puzzle of free indirect discourse”, *Linguistics and Philosophy* 31: 353–395.

Resource texts

- [1] Jen, Gish (1994). “Birthmates”, *Ploughshares* 20 (4): 81-98.
- [2] Johnson, Greg (1996). “The Chinese Box”, *Chicago Review* 42 (2): 1-13. Lee, Harper (1960). *To Kill a Mockingbird*. McIntosh and Otis.

Do direct reference theorists have an attitude problem?

José Manuel Viejo*
Autonomous University of Madrid

i. According to the direct reference account of complex demonstratives (e.g., ‘that man drinking a martini’), the utterance of a simple sentence containing these expressions normally expresses a *singular proposition* —a proposition that contains as constituents the individuals referred to, and not any descriptions or conditions on them. Though there may be some disagreements concerning the details of the proper direct reference story, I think it is fair to say that most philosophers with any view on the matter subscribe to this view (see e.g. Schiffer 1981; Kaplan 1989; Braun 1994, 1995; Recanati 1993; Perry 1997, 2000; Borg 2000; Salmon 2002; and Corazza 2003). This notwithstanding, there have been some important challenges to the direct reference account of complex demonstratives. Prominent among these challenges is Jeffrey King’s (1999, 2001, 2008a, 2008b) objection from propositional attitude ascriptions. I shall examine this objection and suggest that it can be answered.

ii. According to the objection from attitude ascriptions, there are cases where a true propositional attitude ascription does not require the presence of a *singular thought* (i.e. an intentional mental state whose content is appropriately given by a singular proposition), despite containing a complex demonstrative inside its embedded sentence. Consequently, the argument goes, the direct reference account of complex demonstratives must be false. Let me quote King’s complete statement of the objection from attitude ascriptions:

[...] suppose that Greg knows on completely general grounds that exactly one student received a score of one hundred on the exam (e.g., suppose that Greg’s evil but scrupulously honest teaching assistant told Greg this as he tossed Greg his failing effort). Reflecting on the difficulty of the exam, Greg says:

(1) That student who scored one hundred on the exam is a genius.

[...] Suppose that a classmate of Greg’s hears Greg’s teaching assistant tell Greg that exactly one student received one hundred on the exam, overhears Greg’s (sincere) utterance of (1), and on that basis says to another of Greg’s classmates:

* jmanuelviejo@hotmail.com

(2) Greg believes that that student who scored one hundred on the exam is a genius.

[...] The belief ascription seems clearly true in such a case. But how can the direct reference theorist explain this? According to the direct reference theorist, the embedded sentence in (2) expresses a singular proposition that has as a constituent the person the ‘that’ phrase [King’s term for what I am calling ‘complex demonstratives’] in it refers to. So on this view, (2) asserts that Greg stands in the belief relation to this singular proposition. But it seems clear that Greg does not stand in the belief relation to the singular proposition in question. Greg, after all, appears to have only *general beliefs* and has no idea who scored one hundred percent on the examination. Thus it is hard to see how the direct reference theorist can explain the intuition that (2) is true in the situation described. (King 2001, pp. 2-4)

iii. The objection from propositional attitude ascriptions rests on two claims. One of them is the claim that it seems clear that Greg does not stand in the belief relation to a singular proposition containing the individual who scored one hundred on the exam as a constituent. The other is the claim that if the direct reference account of complex demonstratives is true, then (2) asserts that Greg believes a singular proposition about that individual. As a result, direct reference theorists can respond to King’s objection in many different ways. I shall present two lines of response, both of which are consistent with the direct reference theory. The first response, which is the one that I favor, denies King’s claim that if direct reference theories are true, then (2) ascribes to Greg a belief in a singular proposition. More specifically, I shall consider how a certain revision of the hidden indexical theory (Crimmins and Perry 1989, Crimmins 1992) and the translational theory (Richard 1989, 1990) would allow us to reject this claim. The second response explores different ways of arguing (contrary to King) that Greg believes a singular proposition about the individual who scored one hundred on the exam. More specifically, I shall consider how direct reference theories can appeal to semantic instrumentalism (Borg 2007, Kaplan 1989) or cognitivism (Jeshion 2010) to support the claim that Greg stands in a belief relation to a singular proposition about “the genius”. One result of this endeavour is that the direct reference view of complex demonstratives is not clearly refuted neither by facts about ascriptions of propositional attitude nor by intuitions of singular thought.

References

- [1] Borg, E. (2000). Complex demonstratives. *Philosophical Studies*, 9, 229-249. Borg, E. (2007). *Minimal semantics*. Oxford: Oxford University Press
- [2] Braun, D. (1994). Structured characters and complex demonstratives. *Philosophical Studies*, 74, 193-219.

- [3] Braun, D. (1995). What is character? *Journal of Philosophical Logic*, 24, 227-240.
- [4] Corazza, E. (2003). Complex demonstratives qua singular terms. *Erkenntnis*, 59, 263- 283.
- [5] Crimmins, M. (1992). *Talk about beliefs*. Cambridge, Mass.: MIT Press.
- [6] Crimmins, M. and Perry, J. (1989). The prince and the phone booth: Reporting puzzling beliefs. *Journal of Philosophy*, 86, 685-711.
- [7] Jeshion, R. (2010). Singular thought: Acquaintance, semantic instrumentalism, and cognitivism. In R. Jeshion (Ed.), *New essays on singular thought* (pp. 105-140). Oxford: Oxford University Press.
- [8] Kaplan, D. (1989). Demonstratives. In J. Almog, J. Perry, & H. K. Wettstein (Eds.), *Themes from Kaplan* (pp. 481-563). Oxford: Oxford University Press.
- [9] King, J. (1999). Are complex 'that' phrases devices of direct reference?. *Noûs*, 33, 155-182. doi: 10.1111/0029-4624.00148
- [10] King, J. (2001). *Complex demonstratives: A quantificational account*. Cambridge, Mass.: MIT Press.
- [11] King, J. C. (2008a). Complex demonstratives as quantifiers: objections and replies. *Philosophical Studies*, 141, 209-242. doi: 10.1007/s11098-008-9238-9
- [12] King, J. C. (2008b). Complex demonstratives, QI uses, and direct reference. *The Philosophical Review*, 117, 99-117. doi: 10.1215/00318108-2007-026
- [13] Perry, J. (1997). Indexicals and demonstratives. In B. Hale & C. Wright (Eds.), *A companion to the philosophy of language* (pp. 586-612). Oxford: Blackwell.
- [14] Perry, J. (2000). Rip Van Winkle and other characters. In J. Perry (Ed.), *The problem of the essential indexical, expanded edition* (pp. 355-376). Stanford, CA: CSLI Publications.
- [15] Recanati, F. (1993). *Direct reference: From language to thought*. Oxford: Blackwell.
- [16] Richard, M. (1989). How I say what you think. *Midwest Studies in Philosophy*, 14, 317-337.
- [17] Richard, M. (1990). *Propositional attitudes: An essay on thoughts and how we ascribe them*. Cambridge: Cambridge University Press.
- [17] Salmon, N. (2002). Demonstrating and necessity. *Philosophical Review*, 111, 497-537. Schiffer, S. (1981). Indexicals and the theory of meaning. *Synthese*, 57, 43-100.

HISTORIA DE LA CIENCIA

Hacia una Teoría Consciente de la Representación Científica

Milagros Barroso*
Universidad Alberto Hurtado

Propongo el desarrollo de una Teoría Consciente de la Representación Científica, bajo dos premisas centrales: (i) que la relación entre hechos y teorías en ciencia descansa en una imbricación insoslayable entre, por un lado, los modelos teóricos entendidos como realizaciones de la teoría en un sistema, i.e., como las aplicaciones de la teoría en un ámbito particular de la realidad; y por el otro, las teorías como descripciones de una estructura, i.e., como el conjunto de los modelos que definen e identifican los fenómenos y; (ii) que lo anterior puede ser explicado en términos representacionales con base en una concepción de la inducción como método preferente en la práctica científica, que implica lo que el filósofo victoriano William Whewell denominó una Consciliencia de Inducciones, i.e., cuando una inducción contenida en una clase de hechos coincide con una inducción obtenida de otra clase de hechos diferentes (Whewell, 1847).

Con (i) y (ii) hago explícitas dos precisiones teóricas, a saber: (i) que las nociones de “sistema”, “teoría”, “modelo” y “estructura” expuestas en la presente propuesta deben ser entendidas dentro del planteamiento conceptual de la concepción semántica o modelo-teórica de las teorías, según el cual, presentar una teoría es presentar, metalingüística-mente, el conjunto de sus modelos y, (ii) que el tratamiento de la práctica científica que se articula con el análisis al que se hace mención en (i) se circunscribe a la descripción de la ciencia que William Whewell presenta en *The Philosophy of the Inductive Sciences Founded on Their History* (1847), *Novum Organon Renovatum* (1858) y otros textos.

Tanto (i) como (ii) lejos de ser arbitrarias, obedecen a algunas razones de peso: (i) en lo que concierne a la comprensión de las teorías científicas, hay un amplio consenso respecto al progreso que representa la concepción semántica de las teorías en relación a su predecesora, la visión enunciativa de las teorías. Una de las razones la encontramos en el ámbito de la práctica científica: una noción instrumental de las teorías parece conducirse mejor con una noción conceptual de Teoría como un conjunto de modelos en lugar de un conjunto de enunciados. En lugar del análisis micrológico, que hace énfasis en los enunciados y en la relación de deducibilidad; la preferencia por un análisis macrológico aquí propuesta permite, al tomar los modelos como unidad básica de análisis, hacer a un lado dificultades medulares de la concepción enunciativa como la distinción teórico/observacional, el problema de la inducción, entre otros.

* mbelbarroso@gmail.com

Respecto a (ii), Morris Cohen y Ernest Nagel (1990) conciben la elaboración teórica como “una interconexión sistemática de los fenómenos” en la que “se abstraen en general aquellos aspectos que permiten la realización de este objetivo”. En su opinión, cualquier elaboración teórica presupone un proceso de abstracción mediante el cual se busca captar el carácter general de la realidad a través de las relaciones invariables de los fenómenos, con el costo de alejarse con ello del sentido común. Este carácter general se expresa en las teorías científicas a partir de la subsunción de lo particular de los eventos en leyes generales que pretenden explicar teóricamente los fenómenos, las cuales, a su vez, pueden ser explicadas por leyes de mayor generalidad.

Esta subsunción ha sido tradicionalmente entendida como una práctica inductiva apoyada en la teoría estadística o la teoría de probabilidades en sus distintas variantes, así como formando parte de la concepción enunciativa de las teorías. Sin embargo, de un tiempo a esta parte algunos autores (Cobb, 2011; Dethier, 2018) han venido insistiendo en una interpretación según la cual, a partir de la descripción de la ciencia que desarrolla William Whewell en (1847), (1858), entre otros textos, es posible establecer un vínculo entre la práctica inductiva en la ciencia y la concepción semántica o modelo-teórica de las teorías. De ello se sigue, que la inducción no solo es caracterizable desde una concepción enunciativa de las teorías, como es costumbre en Filosofía de la Ciencia, sino que, además, es posible explicar los razonamientos inductivos en la ciencia desde un enfoque semántico o modelo-teórico de las teorías.

La caracterización de las teorías como el conjunto de sus modelos y la proximidad de estos últimos con los fenómenos, lleva, además, a la ineludible cuestión acerca de qué tipo de mediación es la que opera en la noción de modelo, es decir, cuál es la relación entre los modelos y las teorías, por un lado; y los modelos y los fenómenos, por el otro. Una respuesta ha sido dar cuenta de esta relación desde un punto de vista representacional, esto es, sosteniendo que los modelos son, entre otras cosas, construcciones y/o idealizaciones representacionales. Sin embargo, no ha habido acuerdo dentro de la familia semántica respecto a la manera en que estos representan: mientras para los estructuralistas, la representación se entiende en términos de las relaciones matemáticas de similitud, homomorfismo e isomorfismo; otros discrepan de ello, como Bastiaan van Fraassen quien sostiene que los modelos representan puntos o trayectorias en un “espacio de estados”, i.e., un ente matemático en el que se describe las secuencias posibles de valores de las magnitudes en la trayectoria de un objeto.

Las dificultades presentes en las distintas teorías de la representación hasta ahora propuestas han sido documentadas *in extenso* (cfr. Martínez 2001; Frigg 2002; et. al.). Por otra parte, el debate actual a este respecto ha estado dividido entre quienes presentan la representación como un problema ontológico cuya solución requiere dar cuenta de qué tipo de objetos son los modelos y qué es lo que hace que un modelo sea representación de la realidad, con el aditivo de que, además, lo haga desde una perspectiva pluralista (Frigg, 2006); y quienes prescindiendo de consideraciones metafísicas, consideran que

dar cuenta de la noción de representación pasa por hacer explícito cómo esta sucede o, lo que es lo mismo, por señalar cómo una entidad abstracta, i.e., una estructura matemática, puede representar fenómenos concretos (Van Fraassen 2006). Por mi parte, sostengo que la dificultad radica en que no hay una teoría de la representación capaz de articular la representación científica, no solo con los modelos y sus estructuras, sino, además, con la práctica efectiva de los científicos y definiendo que realizar tal articulación es posible desde (i) una postura ontológicamente neutra y (ii) mediante la convergencia del enfoque modelo-teórico de las teorías y la teoría de la consciencia de William Whewell, en la Teoría Consciente de la Representación Científica aquí propuesta.

Bibliografía

- [1] Cobb, A. “History and Scientific Practice in the Construction of an Adequate Philosophy of Science: Revisiting a Whewell/Mill Debate.”, *Studies in History and Philosophy of Science*, 42, 2011:85–93
- [2] Dethier, C. “William Whewell’s Semantic Account of Induction”, *HOPOS: The Journal of the International Society for the History of Philosophy of Science*, vol. 8, 2018: Spring.
- [3] Falguera, J. “La noción de modelo en los análisis de la concepción estructuralista”, *Rev. Ágora: papeles de filosofía*, Vol. 11, 1, 1992: 97-106.
- [4] Frigg, Roman. “Models and Representation. Why structures are not enough”, *Measurement in Physics and Economics Project Discussion Paper Series DP MEAS 25/02*, LEC, 2002.
- [5] Frigg, Roman. “Scientific Representation and the Semantic View of Theory”, *Theoria*, 55, 2006:49-65.
- [6] Martínez, Sergio. “Historia y combinatoria de las representaciones científicas,... ”, *Crítica, Revista Hispanoamericana de Filosofía*, vol. 33, 99, 2001: 75-95.
- [7] Nagel, E. et. al. *Introducción a la lógica y el método científico*. Trad. Néstor Míguez. Tomo II. Buenos Aires: Amorrortu S.A. 1990.
- [8] Van Fraassen, Bas. “Representation: The Problem for Structuralism”, *Philosophy of Science Assoc. 19th Biennial Meeting – PSA 2004*, 2006: 536-547.
- [9] Whewell, William. *The Philosophy of the Inductive Science, The Historical and Philosophical Works of William Whewell*, vols. V-VI, London, Frank Cass & Co. (orig.: 1847).
- [10] Whewell, William. *Novum Organon Renovatum*, London, (orig.: 1858).

Emergent Laws in cosmology

Aldo Filomeno*

Instituto de Filosofía, Pontificia Universidad Católica de Valparaíso

We review the status of some well-known assumptions and limits in the Cosmological Standard Model, assessing whether they support any philosophical accounts of laws of nature. We specifically focus on the support provided to an account of fundamental laws that has for long been studied in various projects in theoretical physics, while it has been neglected in philosophical accounts: a deflationist account of laws as emergent from a non-linear, chaotic, dynamics. If such a dynamics is interpreted as merely describing, it can be seen either as a complement to the Humean account, or as a distinct, underexplored, account of laws, which avoids the main objections of the Humean.

After introducing the features of the FLRW model (and the additions of the λ -CDM), we survey a map in theory-space of possible types of fundamental theories (drawn from [Crowther & Linneman, 2018]). The aim of the talk is to add a neglected option to this map.

1) First, we cite both orthodox and unorthodox projects in cosmology that endorse or assume this type of fundamental theory. Among the mainstream: chaotic (inflationary) cosmology [Misner 69, Linde 83], and among the unorthodox: the ‘Random dynamics’ project [Nielsen et. al. 91], and entropic gravity projects [Jacobson 95, Verlinde 17].

2) We assess how this type of theory fares regarding the traditional non-empirical criteria that guide the search for a fundamental theory. Drawing from a list by [Crowther 2019], we focus on the widely debated criteria of simplicity, unification, and naturalness. A fundamental chaotic dynamics does not seem to fulfill them. One option, increasing in the literature, is to neglect these criteria. Yet, even if we take them seriously, it can be argued that the type of theory under consideration would wash out the worries of naturalness and simplicity at the fundamental level.

3) Interpreting the fundamental dynamical complexity without an ontological commitment to “guiding” laws might reasonably raise suspicion. There are accounts of emergent laws, but they emerge from other laws (for instance, effective field theories from other theories at higher energy scales). To address this, we cite the study by [Filomeno 2019], and references therein (basically, well-known results in dynamical systems theory), which aim to assess the plausibility of the emergence of laws, not from a deterministic chaos, but from a truly process-random state-space trajectory. Such proposals rely on

* aldofilomeno@gmail.com

the postulation of allegedly non-dynamical conditions, such as kinematical or boundary conditions.

4) This new option in the map is compatible with the usual implicit view of the whole universe being described by a FLRW model. Still, this option is mutually reinforced by the alternatives to the FLRW model currently investigated in theoretical physics. To show this, we begin by citing the epistemological limits of the cosmological standard model, as it is widely acknowledged in textbooks (e.g. [Dodelson 2021, Hartle 2003], cf. [Beisbart 06, 09, Uzan 09]). The simplest and strongest point in support of alternatives to FLRW models is the lack of epistemic warrant to extrapolate our local model to the global structure (whose size could e.g. be dozens of orders of magnitude larger). (Additionally, (controversial) fine-tuning problems, such as the detected accelerated expansion of the observable universe, related to λ , have been interpreted as support for a universe vastly larger, and different, from the observable region.). This picture can be framed together with the new option in the map, whereby the aforementioned boundary conditions are parochial contingencies of a local region. For instance, following [Smeenk 2013, 644], if we accept that something like the inflationary phase transition at the early universe occurred, one option is that the inflaton field tunneled into a local minima with certain properties.

One virtue of this type of view is that it answers the question of “why these laws rather than others”. There is no need to invoke a multiverse: the answer relies on the contingency of the initial and boundary conditions, and an explanation for the emergence of effective laws at lower-energy levels. As opposed to the Best System account, this deflationist way of understanding laws explains both non-accidental regularities and their ubiquity (the extremely high degree of regularity of the Humean mosaic).

References (partial selection)

- [1] Barrow, J. D. (1977). A chaotic cosmology, *Nature* 267, 117-120.
- [2] Batterman, Robert W. (2001). *The devil in the details: Asymptotic reasoning in explanation, reduction, and emergence*, Oxford University Press.
- [3] Dodelson, S. (2003). *Modern Cosmology* (Fermilab and Chicago U.) Academic Pr.
- [4] Beisbart, C. (2016). Philosophy and Cosmology. In Paul Humphreys (ed.), *Oxford Handbook in the Philosophy of Science*. Oxford, Vereinigtes Königreich: pp. 817-835.
- [5] Beisbart, C. (2009). Can We Justifiably Assume the Cosmological Principle in Order to Break Model Underdetermination in Cosmology? *Journal for General Philosophy of Science* 40 (2): 175-205.

- [6] Chkareuli, J.L., Froggatt, C.D. and Nielsen, H.B. (2011). Spontaneously Generated Tensor Field Gravity, *Nucl.Phys.* B848, 498-522.
- [7] Chadha, S. and Holger Bech Nielsen (1983). Lorentz invariance as a low energy phenomenon, *Nuclear Physics B* 217, no. 1, 125-144.
- [8] Crowther, K. (2019). *When do we stop digging? conditions on a fundamental theory of physics, What is 'Fundamental'?*, Springer.
- [9] Ellis, George F.R. (2014). Issues in the philosophy of cosmology. *Studies in History and Philosophy of Modern Physics*.
- [10] Filomeno, A. (2019). Stable regularities without governing laws?, *Studies in History and Philosophy of Modern Physics* 66, 186-197.
- [11] Froggatt, C.D. and Nielsen, H.B. (1991). Origin of symmetries, *World Scientific*.
- [12] Jacobson, T. (1995). Thermodynamics of space-time: The Einstein equation of state, *Phys.Rev.Lett.* 75, 1260-1263.
- [13] Linde, A.D. (1983). Chaotic inflation, *Physics Letters B* 129, n° 3, 177-181.
- [14] Misner, Charles W. (1969). Mixmaster universe, *Phys. Rev. Lett.* 22, 1071-1074.
- [15] Mukohyama, Shinji and Uzan, Jean-Philippe (2013). From configuration to dynamics: Emergence of Lorentz signature in classical field theory, *Phys.Rev.* D87, n° 6, 065020.
- [16] Peebles, P. J. E. (2021). *Cosmology's Century: An Inside History of Our Modern Understanding of the Universe*, Princeton University Press.
- [17] Smeenk, Chris (2013). Philosophy of Cosmology. In Robert Batterman (ed.), *Oxford Handbook of Philosophy of Physics*. Oxford: Oxford University Press. pp. 607-652.

How early humans made science possible

Alexander Krauss*

London School of Economics / University of Barcelona

One of the most foundational questions of science is how we developed science. The common view in science, philosophy of science and history of science is that science emerged largely around the 17th century in Europe (e.g. Zilsel 2003; Eisenstein 1980; Henry 2008; Logan 1986). That common view of science has been maintained because, in studying the question of the origin and foundations of science, researchers have focused commonly on proximate factors (or causes) as explanations, such as the spread of capitalism, Christian values and the printing press that are viewed as giving way to the work of scholars like Bacon, Galileo and Newton (Zilsel 2003; Merton 1938; Eisenstein 1980).

Some scientists and philosophers have looked deeper than the proximate factors since the 17th century by attempting to identify the underlying foundational factors that have allowed us to develop science in the first place. Leading cognitive scientist Dennett (1991) argues that a massive reprogramming of the mind made science possible. Leading evolutionary psychologists Pinker (2010) and Cosmides and Tooby (2013) have stressed features of the mind such as our ability to use our folk physics to deal with the world around us. Leading evolutionary biologist Laland (2017) and anthropologist Henrich (2016) stress how our cognitive capacity for culture has been the central driver. Leading developmental psychologist Tomasello (2019, 1999) similarly emphasises the function of social cooperation and coordination.

These researchers across different fields focus mainly on what broad factor can explain the development of science. Existing studies do not yet offer an integrated explanation of the origin of science that assesses particular human abilities. To better address the question, the paper integrates evidence from evolutionary biology, cognitive sciences, methodology, archaeology and philosophy of science. The paper thereby provides a methodological account of the fundamental factors that have helped make science possible. It outlines in particular our biological and cognitive capacities for systematic observation, problem-solving, basic experimentation etc. (which we largely developed before the Paleolithic) and complex cognitive and social capacities including more complex language, tools and imagination (which we largely developed during the Paleolithic). Our early ancestors used trial and error to improve the tools and technologies they created (Mithen 2002; Toth et al. 1993) and they accumulated experimental knowledge about nutrition

* a.krauss@lse.ac.uk

and botany (Atran and Medin 2008; Blurton-Jones and Konner 1976). Science did not just emerge but gradually developed over time. The paper thereby aims to provide a new framework for understanding the history of science. It concludes drawing implications on science and our understanding of the scientific method.

Selected references

- [1] Atran, Scott; Douglas Medin. (2008). Introduction. In: *The Native Mind and the Cultural Construction of Nature*. Cambridge, Massachusetts: MIT Press.
- [2] Blurton-Jones, N.; Konner, M. (1976). Kung knowledge of animal behaviour. In, R. Lee and I. DeVore (eds.), *Kalahari Hunter-Gatherers*. Cambridge University Press.
- [3] Carruthers, P. (2002). The roots of scientific reasoning: infancy, modularity and the art of tracking. In: *The Cognitive Basis of Science*. Cambridge University Press.
- [4] Cosmides, L; J. Tooby. (2013). Evolutionary psychology: new perspectives on cognition and motivation. *Annu Rev Psychol*, 64, pp. 201-229.
- [5] Dennett, D. (1991). *Consciousness Explained*. London: Allen Lane.
- [6] Eisenstein, E. (1980). *The Printing Press as an Agent of Change*. Cambridge: Cambridge University Press.
- [7] Giere, R. (1987). The cognitive study of science. In N. Nersessian (Ed.), *The process of science*. Springer.
- [8] Henrich, Joseph. (2016). *The Secret of Our Success: How Culture is Driving Human Evolution, Domesticating our Species, and Making us Smarter*. Princeton University Press.
- [9] Henry, John. (2008). *The Scientific Revolution and the Origins of Modern Science*. Houndsmill: Palgrave Macmillan, 3rd ed.
- [10] Laland, Kevin. (2017). *Darwin's Unfinished Symphony: How Culture Made the Human Mind*. Princeton University Press.
- [11] Logan, R. (1986). *The alphabet effect*. New York: Morrow.
- [12] Merton, R. (1938). Science, Technology and Society in Seventeenth Century England. *Osiris*, 4, 360-632.
- [13] Mithen, S. (2002). Human evolution and the cognitive basis of science. In: *The Cognitive Basis of Science*. Cambridge University Press.

- [14] Pinker, S. (2010). The cognitive niche: coevolution of intelligence, sociality, and language. *Proc Natl Acad Sci*, 107, pp. 8993-8999.
- [15] Tomasello, Michael. (1999). *The cultural origins of human cognition*. Cambridge: Harvard University Press.
- [16] Tomasello, Michael. (2019). *Becoming Human: A Theory of Ontogeny*. Harvard University Press.
- [17] Toth, N. et al. (1993). Pan the tool maker: investigations into the stone tool-making and tool-using capabilities of a bonobo (*Pan paniscus*). *Journal of Archaeological Science*.
- [18] Zilsel, Edgar. (2003). *The Social Origins of Modern Science*. (Eds. Raven, Krohn and Cohen); Dordrecht, Netherlands: Kluwer Academic Publishers.
- [19] Zvelebil, M. (1984). *Clues to recent human evolution from specialized technology*. *Nature*.

Transferencia conceptual y el rol de la metáfora en ciencia: el caso de la noción de ‘división del trabajo’ en la biología del siglo XIX

Daniel Labrador Montero*
University of Salamanca

La presente comunicación tiene como objetivo analizar un caso concreto acerca de la transferencia conceptual entre la economía política y la biología producida en el siglo XIX. El estudio de caso será el traslado del concepto metafórico de ‘división del trabajo’ desde las ciencias sociales y la economía política a la biología (Limoges, 1994; D’hombres, 2012, 2015).

La exposición comenzará estableciendo el marco teórico del que se parte en lo relativo al papel de la metáfora en la ciencia. En este punto, será esencial destacar la teoría de Maasen y Weingart (2003), que es de especial relevancia para entender el rol de la metáfora de la división del trabajo en la teoría biológica. Estos autores señalan dos fases diferenciadas en la transferencia conceptual interdisciplinar. Por un lado, estaría la fase de transferencia. En esta etapa las metáforas transferidas de una teoría a otra juegan un papel fundamentalmente heurístico que puede llegar a permitir una “reorganización de los fenómenos” o una “nueva manera de problematizar” el asunto. Así, si el traslado conceptual es exitoso, la metáfora es capaz de ilustrar nuevas conexiones que antes no se habían tenido en cuenta (Maasen, 1994, pp. 22-23). Aun así, en este tramo, los conceptos importados siguen apreciándose como intrusos o forasteros, como metáforas.

Una segunda fase sería la de transformación. Aquí se hace referencia a la asimilación o procesamiento por parte de la teoría importadora de los conceptos que en un primer momento resultaban ajenos. Ese paso no se da de forma necesaria, es decir, es posible que la metáfora se quede en su función heurística y jamás pase a formar parte del corpus de la teoría de importación. No obstante, en otros casos la teoría asume la metáfora y la literaliza (la convierte en familiar, en concepto propio).

Tras haber realizado esta introducción, será oportuno hacer un breve repaso de las implicaciones de la noción de división del trabajo en la economía política de la época. Para ello, se explicará sucintamente las diferencias básicas entre la división social, técnica y territorial del trabajo. Esta distinción se hará en aras de entender cuáles de esas categorías son los que se trasladan a la biología.

* danilabra68@gmail.com

Una vez aclarado esto, se acometerá el análisis del primer caso en el que se produjo la transferencia del concepto a la biología. Esta importación fue llevada a cabo por el fisiólogo Henri Milne-Edwards. Así, este autor inventó la noción de división fisiológica del trabajo, que hace referencia a la especialización de los órganos para dedicarse a una función específica. Al introducir este concepto en la fisiología, lo que buscaba era encontrar un criterio de complejidad para la jerarquización de los seres vivos conforme a su complejidad. Por tanto, Milne-Edwards acudió a un concepto extradisciplinario para ir dónde las herramientas disciplinarias no se lo permitían, he hizo de la división fisiológica del trabajo el criterio de complejidad: a mayor división del trabajo mayor complejidad del organismo.

De esta manera, hay una transferencia desde el contexto económico al biológico, y el propio Milne-Edwards admite la procedencia del concepto, a saber, los economistas políticos. Ahora bien, un análisis profundo del concepto nos lleva a obtener la siguiente conclusión: el concepto al que acude el francés es al de división técnica del trabajo (es decir, dentro de una manufactura).

Por otra parte, Milne Edwards no es el único en utilizar la noción de división del trabajo en su teoría biológica. Charles Darwin hace lo propio, lo cual para muchos es prueba del gran influjo que hubo en el naturalista inglés por parte de los economistas (por ejemplo, Schweber, 1980, 1985; Gould, 2002; Hayek, 1967, 1988, 1991). Sin embargo, no tenemos evidencia real acerca de dónde pudo obtener Darwin algunas nociones económicas respecto a la división del trabajo. Teniendo en cuenta cuál era su contexto intelectual y político, así como sus lecturas, tales nociones pudieron venir de muchas y diversas fuentes.

De lo que sí tenemos evidencia es de que Darwin asume la división fisiológica del trabajo y cita a Milne-Edwards. El naturalista británico considera (Darwin, 1872, p. 89) que lo está haciendo es extender tal noción a la naturaleza en su totalidad y, por tanto, dando lugar así a lo que se podría tildar de división ecológica del trabajo. Ya no hablamos de la división del trabajo entre órganos dentro de un organismo, sino entre organismos dentro de la economía natural (hoy diríamos en un ecosistema).

Darwin hace de tal noción un elemento fundamental de la comprensión de su principio de divergencia, la otra piedra angular de su teoría junto a la selección natural (Mayr, 1992; Tammone, 1995; Kohn, 2009; Pearce, 2010; Richards, 2012). No obstante, hay diferencias radicales entre la metáfora de Darwin y la de Milne-Edwards. Las pruebas textuales, como se mostrará en la exposición, parecen indicar que el concepto de división ecológica del trabajo se asemeja más al concepto de división social que de división técnica del trabajo.

Así las cosas, la principal conclusión que se esgrimirá de dicho examen es que, a diferencia del caso de Milne-Edwards, cuyo uso del concepto se encuentra en la fase de transferencia, en el caso del concepto darwiniano se está iniciando el proceso de transformación. Por tanto, el concepto importado va más allá de la simple labor heurística y cumple un papel interno y esencial dentro de la teoría.

Se trata de un concepto transferido en cuanto se tiene en cuenta la teoría de la división social del trabajo sobre el que se apoya la metáfora. Sin embargo, a la vez, es una transformación radical del concepto ya usado por Milne-Edwards en su teoría fisiológica. En suma, se podría decir, que el concepto darwiniano se encuentra en la frontera entre las fases de transferencia y transformación.

Bibliografía citada

- [1] Darwin, C. (1872). *The Origin of Species by Means of Natural Selection or the Preservation of Favoured Races in the Struggle for Life*. 6th Edition. London: John Murray, Albemarle Street.
- [2] D’Hombres, E. (2012). The ‘Division of Physiological Labour’: The Birth, Life and Death of a Concept. *Journal of the History of Biology*, 45, pp. 3-31.
- [3] D’Hombres, E. (2015). The Darwinian Muddle on the Division of Labour: An Attempt at Clarification. *History and Philosophy of the Life Sciences*, 38, pp. 1-22.
- [4] Gould, S. J. (2002). *The Structure of Evolutionary Theory*. Cambridge, Massachusetts y Londres: The Belknap Press of Harvard University Press.
- [5] Hayek, F. A. (1967). *Studies in Philosophy, Politics, economics, and History of Ideas*. Londres: Routledge and Kegan Paul.
- [6] Hayek, F. A. (1988). The Fatal Conceit: The Errors of Socialism. En *Collected Works of F. A. Hayek* (vol. 1): Londres: Routledge.
- [7] Hayek, F. A. (1991). The Trend of Economic Thinking: Essays on Political Economist and Economics History. En *Collected Works of F. A. Hayek* (vol. 3): Londres: Routledge.
- [8] Kohn, D. (1985). Darwin’s Principle of Divergence as Internal Dialogue. En: Kohn, D. (ed.) *Darwinian Heritage*. New Jersey: Princeton University Press, pp. 245-258.
- [9] Kohn, D. (2009). Darwin’s Keystone: The Principle of Divergence. En Ruse, M. y Richards, R. J. (eds.) *The Cambridge Companion to the Origin of Species*. Cambridge: Cambridge University Press, pp. 87-108.
- [10] Limoges, C. (1994). Milne-Edwards, Darwin, Durkheim and the Division of Labour: A Case Study in Reciprocal Conceptual Exchanges Between the Social and the Natural Sciences. En Cohen, T. B. (ed.), *The Natural Sciences and the Social Sciences*. Kluwer Academic Publishers, pp. 317-343.

- [11] Maasen, Sabine (1994). Who is Afraid of Metaphors? En Sabine Maasen, Everett Mendelsohn y Peter Weingart (eds.), *Biology as Society, Society as Biology: Metaphors* (11-36), Dordrecht: Kluwer Academic Publisher.
- [12] Maasen, Sabine y Weingart, Peter (2003). *Metaphors and the Dinamic of Knowledge*. London: Routledge.
- [13] Mayr, E. (1992). Darwin's Principle of Divergence. *Journal of the History of Biology*, 25 (3), pp. 343-359.
- [14] Pearce, T. (2010). "A Great Complication of Circumstances". Darwin and the Economy of Nature. *Journal of the History of Biology*, (43), pp. 493-528.
- [15] Richards, R. J. (2012). Darwin's principles of divergence and natural selection: Why Fodor was almost right. *Studies in History and Philosophy of Biological and Biomedical Sciences* 43(1), pp. 256-268.
- [16] Schweber, S. (1980). Darwin and the Political Economist: Divergence of Character. *Journal of the History of Biology* 13 (2), pp. 195-289.
- [17] Schweber, S. (1985). The wider British context in Darwin's theorizing. En Kohn, D. (ed.), *The Darwinian heritage*, New Jersey: Priceton University Press, pp. 35-69.
- [18] Tammone, W. (1995). Competition, the Division of Labor, and Darwin's Principle of Divergence. *Journal of the History of Biology*, 28 (1), pp. 109-131.

La (re)emergencia de las filosofías pluralistas de la ciencia: variaciones divergentes sobre un tema

Joseba Pascual Alba*
Universidad del País Vasco (UPV-EHU)

Resumen

En los comienzos de la Filosofía de la Ciencia como disciplina, era habitual hablar de una Ciencia Unificada, opuesta a la pluralidad. Sin embargo, tanto otras tradiciones como los sucesores del neopositivismo no vieron vicio alguno en el pluralismo, sino una realidad irrenunciable. Así, llegamos a los nuevos teóricos del pluralismo científico: J. Dupré, N. Cartwright, I. Hacking y la “Epistemología Histórica” (H.J. Rheinberger, L. Daston) abrieron nuevos mapas de esta pluralidad; más tarde, otros como H. Longino, H. Chang, M. Massimi, S. Ruphy o S.D. Mitchell han renovado la problemática pluralista. El problema es que cada pluralismo tiene su especificidad y abarca tan solo una parte de los problemas filosófico-científicos. Así, el objetivo de esta comunicación es realizar un análisis comparativo de las filosofías pluralistas de la ciencia, con el objetivo taxonómico-filosófico de discernir y aclarar dichas posiciones, para una posible ulterior propuesta más completa.

Propuesta de comunicación

La Filosofía de la Ciencia ha sido, en cierto modo, una disciplina de tradición *unitarista*. Esto es, desde sus inicios como disciplina académica se ha tendido a entender la investigación científica como algo a entender de manera unívoca: el estudio de La Ciencia. Filósofos como E. Mach o sus seguidores neopositivistas –R. Carnap, O. Neurath– reivindicaron la *Ciencia Unificada* y, más tarde, reduccionista (Oppenheim y Putnam, 1958). Sin embargo, todos coinciden en algún punto de su *diagnóstico* o descripción pluralista de las ciencias –esto es, que, *de facto*, hay muchas ciencias y no una única ciencia. Su unitarismo vendrá dado porque, de un modo u otro, consideran el pluralismo como *pasajero* –o *contingente*.

Sin embargo, otros filósofos de la ciencia, de manera más o menos explícita, aceptaron el *diagnóstico* pluralista de las ciencias, pero reaccionaron contra la aplicación del *remedio* unitarista –por seguir con la analogía médica. Tal es el caso de la saga de las revoluciones científicas (Kuhn, Lakatos, Laudan). También llegó el relativismo de P. Feyerabend (1975) y su “*anything goes!*”, que aún influyen en diversos movimientos

* joseba.pascual.alba@gmail.com

“anti-ciencia” –“anti-vacunas”, negacionistas del Cambio Climático y, actualmente, del “Covid-19”.

En este sentido, desde entonces, se han venido desarrollando diversas respuestas al unitarismo. Es un hecho que en Filosofía de la Ciencia se viene hablando cada vez más sobre el pluralismo científico (Fodor, 1974; Galison & Stump, 1996; Kellert, Longino y Waters, 2006; Kusch, 2017; Vagelli et al., 2021). El primer movimiento de impacto real lo tuvieron los miembros de la “Stanford School”: J. Dupré (1993; 2018) construye su pluralismo a la contra del reduccionismo físico-químico en las ciencias de la vida. Este conecta, además, su tesis epistémica con la ontología: el pluralismo y la irreductibilidad entre ciencias se deben al pluralismo inherente a la realidad: una realidad plural, y “desordenada”. Así, Dupré defiende un “realismo promiscuo”. Para N. Cartwright (1999), el mundo mismo está *moteado* (“dappled”), no sometido a un orden legal de estructuras rígidas (pluralismo causal). Por eso, niega la posibilidad del reduccionismo ya que las ciencias constituirían un conjunto fragmentario de leyes diversas, puesto que la propia realidad está estructurada en diferentes dominios y sistemas de leyes que no se relacionan entre sí de manera uniforme (Cartwright, 1983, 1999). I. Hacking (1983, 1996) plantea seis “estilos de razonamiento científico”. Cada campo científico tendría su propio estilo de razonamiento, mediante el cual le es posible trabajar y diferenciarse del resto de ciencias, cada una con sus leyes, experimentos, modalidades, etc., construyendo una diferenciación no sólo metodológica, sino también ontológica –“realismo de entidades”. Familiar con las anteriores, en el ámbito de la llamada “Epistemología Histórica” (G. Bachelard, L. Fleck) el pluralismo se introduce en el estudio de las ciencias partiendo de la historicidad de objetos, conceptos o campos científicos (H.J. Rheinberger, 1997, 2007; Daston, 2000).

Por su parte, H. Chang (2012) defiende el *pluralismo experimental* al que obliga el estudio histórico-filosófico de las ciencias. Se lanza a promover un “pluralismo epistémico activo y normativo”, o sea, una filosofía de la ciencia con el fin de “promover la pluralidad para cosechar sus beneficios” (2012: 268). Pero, no todo vale (*vs* Feyerabend), aunque sí valen muchas cosas –“many things go!”. Define así el pluralismo en ciencia como “la doctrina que aboga por el cultivo de múltiples sistemas de práctica en cualquier campo dado de la ciencia” (2012: 260) y defiende una suerte de “presentismo pluralista” en Historia de la Ciencia (Chang, 2020). Simpatiza con cierto pluralismo ontológico y defiende el anti-reduccionismo. Se define como realista pragmático, en cierto modo, y reivindica que cualquier “realista real” –o verdadero– es, en realidad, pluralista (Chang, 2017) (*Cf.* asimismo Chakravartty, 2011).

Otra de las propuestas más importantes la ha realizado S. Ruphy (2013), mediante lo que denomina “foliated pluralism”. Recogiendo el testigo de los “stanfordianos”, realiza una labor de integración entre los “estilos” de Hacking –que suponen la aparición en la realidad de *nuevas entidades*– con un “enriquecimiento ontológico” –que amplía y diversifica las clases de proposiciones científicas, que nunca son finales ni completas. Su

pluralismo tiene las siguientes propiedades: *transdisciplinariedad* (mismos *estilos* superpuestos en diversas disciplinas); *no-exclusividad* y *sincronicidad* (combinación simultánea de diversos *estilos*); y *cumulatividad* (los *estilos* se acumulan en el desarrollo de las ciencias).

Por otra parte, R. Giere (2006) ha propuesto un “pluralismo perspectivista”, como consecuencia de su visión perspectivista de las ciencias. Más recientemente, M. Massimi (2017; Massimi y McCoy, 2020) viene ofreciendo una versión renovada de este perspectivismo. Su filosofía trata de coordinar cierto realismo científico con un “pluralismo perspectivista e integrador”, a la contra tanto del monismo como del relativismo y de los movimientos “anti-ciencia” mencionados. Se distingue, asimismo, de pluralismos epistémicos (H. Longino) y “explicativos” (C. Fehr).

Por último, cabe destacar también el “pluralismo integrador” de S. D. Mitchell (2002), quien trata de ofrecer un nuevo modelo para entender el pluralismo científico. Mitchell (1992) realiza una primera distinción entre “pluralismos competitivos” y “pluralismos compatibilistas”. Sus investigaciones sobre distintos casos biológicos, le sirven de base para argumentar a favor de la *integración* de los diversos modelos en el terreno de las explicaciones científicas: el pluralismo es necesario a un nivel teórico en las ciencias, lo que no obsta para la coexistencia con la integración en el nivel de la explicación concreta –solo hay una “historia causal”. El pluralismo integrador de Mitchell defiende una suerte de realismo pragmático *sui generis*: las representaciones científicas son parciales –adaptadas para cumplir con objetivos explicativos dados de manera pragmática. Es por eso que coexisten una pluralidad de métodos, modelos y representaciones (Mitchell y Gronenborn, 2017).

Como vemos, *el pluralismo se dice de muchas maneras*. El objetivo de esta comunicación, es realizar un análisis comparativo, una taxonomía y una labor de criba de las diversas filosofías pluralistas. Un *mapeo* de los diversos pluralismos nos advierte de que: a) no existe una clasificación exhaustiva de los pluralismos; b) las meras taxonomías no son obvias, puesto que la multiplicidad de criterios y sus cruces aumenta con cada pluralismo; c) la noción de “pluralismo” es vaga –o plurívoca– y necesita de clarificación; y d) un mapa exhaustivo sobre los pluralismos nos permitirá explicar mejor su emergencia actual. Asimismo, esta tarea servirá para proponer un pluralismo más ajustado a la investigación científica real.

Referencias

- [1] Cartwright, N. (1983). *How the Laws of the Physics Lie*. New York: Basic Books.
- [2] Cartwright, N. (1999). *The Dappled World: A Study of the Boundaries of Science*. Cambridge: Cambridge University Press.

- [3] Chakravartty, A. (2011). Scientific realism and ontological relativity. *The Monist* 94, 157-180 Chang, H. (2012). *Is Water H₂O? Evidence, Realism and Pluralism*. London: Routledge.
- [4] Chang, H. (2017). Is Pluralism Compatible with Scientific Realism? En Staasi, J. (ed.) (2017). *The Routledge Handbook of Scientific Realism*. London: Routledge, pp. 176-186.
- [5] Chang, H. (2020). Presentist History for Pluralist Science. *J Gen Philos Sci* (2020) Daston, L. (ed.). (2000). *Biographies of Scientific Objects*. Chicago: UCP.
- [6] Dupré, J. (1993). *The Disorder of Things: Metaphysical Foundations of the Disunity of Science*. Cambridge: Harvard University Press.
- [7] Dupré, J. (2018). Processes, organisms, kinds, and the inevitability of pluralism. *Individuation, Process, and Scientific Practices*, 21-39.
- [8] Feyerabend, P. K. (1975). *Tratado contra el método*. Madrid: Tecnos, 1986.
- [9] Fodor, J. A. (1974). Special sciences (or: The disunity of science as a working hypothesis). *Synthese*, 28 (2):97-115.
- [10] Giere, R. (2006) *Scientific Perspectivism*. Chicago: University of Chicago Press.
- [11] González, W. (2020). *Methodological Prospects for Scientific Research. From pragmatism to pluralism*. Dordrecht: Springer.
- [12] Hacking, I. (1983). *Representar e intervenir*. Barcelona: Paidós, 1996.
- [13] Hacking, I. (1996). The Disunities of the Science. En P. Galison & D. Stump (eds.) pp. 37-74. Kellert, S., Longino, H. & Waters, C. (eds.) (2006). *Scientific Pluralism*. Minneapolis: University of Minnesota Press.
- [14] Kusch, M. (2017). Epistemic relativism, scepticism, pluralism. *Synthese* 194 (12):4687-4703. Massimi, M. (2017). Perspectivism. In J. Saatsi (ed.), *The Routledge Handbook of Scientific Realism* (pp. 164-175).
- [15] Massimi y McCoy (eds.) (2020). *Understanding Perspectivism*. London: Routledge.
- [16] Mitchell, S.D.: 1992, 'On Pluralism and Competition in Evolutionary Explanations', *American Zoologist* 32, 135-144.
- [17] Mitchell, S. D. (2002). Integrative pluralism. *Biology and Philosophy* 17 (1):55-70. Mitchell, S. D. (2003). *Biological Complexity and Integrative Pluralism*. CUP: Cambridge.

- [18] Mitchell, S.D. (2020). Perspectives, Representation, and Integration. In Massimi y McCoy (eds.) (2020). *Understanding Perspectivism*. London: Routledge.
- [19] Mitchell, S.D. y Gronenborn, A. (2017). After Fifty Years, Why Are Protein X-ray Crystallographers Still in Business? *Brit. J. Phil. Sci.* 68, 703-723.
- [20] Oppenheim, P. & Putnam, H. (1958). Unity of Science as a Working Hypothesis. *Minnesota Studies in the Philosophy of Science* 2:3-36.
- [21] Rheinberger, H. J. (1997). *Toward a History of Epistemic Things*. Stanford: Stanford University Press.
- [22] Rheinberger, H. J. (2007). *On historizing epistemology*. Stanford: Stanford University Press, 2010. Ruphy, S. (2013). *Scientific Pluralism Reconsidered: A New Approach to the (Dis)Unity of Science*. Pittsburgh: University of Pittsburgh Press, 2016.
- [23] Vagelli, M., Loison, L. & Moya-Diez, I. (2021). Thinking Crossroads: from Scientific Pluralism to Pluralist History of Science. *J Gen Philos Sci* (2021).

Sobre el quiebre del aparato conceptual kuhniano en el terreno de la biología evolutiva contemporánea

M.^a de Los Ángeles Pérez del Amo*
Universidad Complutense de Madrid

Analizaremos el aparato conceptual presentado y defendido por Kuhn resaltando las carencias que este modelo tendría para hacerse cargo del desarrollo histórico de la biología, así como de los debates y las cuestiones propias de la biología evolutiva más actual. De este modo, traeremos a colación las nociones de paradigma y ciencia normal, introduciendo el modelo histórico defendido por el mismo –con los sucesivos periodos de crisis e instauración de nuevos paradigmas y la supuesta inconmensurabilidad entre estos–. Estas herramientas metodológicas se pondrán a prueba con el caso particular de una disciplina como la biología, apoyándonos en el artículo de Alejandro Fábregas-Tejeda y Francisco Vergara Silva denominado: *The emerging structure of the Extended Evolutionary Synthesis: where does Evo-Devo fit in?* Donde se muestran las carencias de la perspectiva kuhniana, presentando, al mismo tiempo, un nuevo punto de partida desde la así denominada Síntesis Extendida.

Propuesta de Comunicación

En esta comunicación nos disponemos a estudiar el aparato conceptual que Kuhn expone a lo largo de su escrito denominado *La estructura de las revoluciones científicas*, del mismo modo, traeremos a colación algunos matices introducidos por este mismo autor en escritos posteriores a esta publicación, como en *La tensión esencial*. De este modo, a través de su novedosa concepción de la historia de la ciencia, iremos introduciendo las que localizaríamos como nociones fundamentales, a saber: la noción de *paradigma* –aludiendo a esa fase de consenso característica del periodo de normatividad científica–, la noción de *ciencia normal* –recurriendo a la explicación clásica de Kuhn, caracterizando este periodo y las actividades realizadas en él con un alto grado de determinismo–, así como el modelo histórico general que defiende –aludiendo a la profunda inconmensurabilidad entre los diversos paradigmas que en un primer momento sostiene–. Todo esto nos servirá de base para, posteriormente, analizar en profundidad el desarrollo histórico y conceptual que se ha venido dando en disciplinas científicas que se desmarcan de la evolución de la Física, en este caso, centraremos nuestra mirada en la biología y el debate

* maripe32@ucm.es

actual sostenido en este sentido. Para lograr esto recurriremos al artículo de Alejandro Fábregas-Tejeda y Francisco Vergara Silva denominado: *The emerging structure of the Extended Evolutionary Synthesis: where does Evo-Devo fit in?*, donde nos presentan los diferentes modelos que se han ido defendiendo y las disputas generadas, especialmente en lo relativo al mantenimiento de las nociones kuhnianas para explicar la evolución histórica de los modelos y las teorías biológicas.

Siguiendo con la estela kuhniana, examinaremos especialmente el modelo basado en la unicidad del pensamiento evolutivo propuesto por Pigliucci, el cual, utilizando el aparato conceptual de Kuhn, trata de adaptar tales nociones al proceso evolutivo propio de las teorías biológicas, tratando, al mismo tiempo, de compatibilizar tal modelo con los debates contemporáneos y las novedades generadas. El autor defendería la idea de un gran paradigma que abordaría todas las teorías y propuestas generadas desde el ámbito de la biología, partiendo de la idea de que no ha existido un cambio lo suficientemente grande en la historia de la biología que podamos datar verdaderamente como un cambio de paradigma. De este modo, defiende que la biología no ha experimentado cambios sustanciales desde los trabajos de Darwin y Wallace, manteniendo desde ellos un mismo marco conceptual, pero asumiendo diversas ampliaciones, partiendo en todo momento de ese núcleo inamovible y unificador. En este sentido, elaboraremos una importante crítica al modelo propuesto, basándonos en las variaciones del marco conceptual kuhniano que se ha visto obligado a realizar para acoplarlo al marco propio de la biología. De esta forma, centraremos nuestra crítica al aparato conceptual kuhniano partiendo de la complejidad y la pluralidad que caracteriza los estudios propios del ámbito de la biología, mostrando la imposibilidad de reducir todo este múltiple desarrollo a un esquema conceptual que no ha sido pensado para este tipo de saberes –refiriéndonos a la evasiva kuhniana con respecto a la biología y a la centralidad de su estudio, refiriéndose en todo momento a modelos propios de la física cuando habla de ciencia–.

Tras analizar y evaluar el modelo kuhniano propuesto por Pigliucci, se introducirá a grandes rasgos un nuevo punto de partida alternativo, capaz de hacerse cargo de esta pluralidad, así como de los debates actuales y las diferentes disciplinas dentro de lo que abarcaría la gran etiqueta de la Biología. Comenzamos, así, con la exposición de los preceptos generales de la nueva Síntesis Extendida y el complejo modelo que desde ella defenderíamos. Desde esta nueva perspectiva se podría mantener esa pluralidad o diversidad de teorías y discursos, pero sin la necesidad de mantener la centralidad conceptual que desde los modelos previos venía impuesta por las extensiones de un mismo marco conceptual basado en las teorías de Darwin y Wallace. De esta forma, el nuevo modelo se fundamentaría en la imagen de una difusa red de diversos modelos y representaciones vinculadas con diferentes prácticas y contextos, permitiendo una gran variedad de puntos cardinales desde los cuales nos localizaríamos para abordar la multiplicidad de problemáticas que desde una disciplina como la biología somos capaces de comprender, evitando,

al mismo tiempo, la necesidad de un centro estático desde el cual derivan el resto de teorías y propuestas normativas.

Bibliografía

- [1] Fábregas-Tejeda, A., & Vergara-Silva, F., (2018). The emerging structure of the Extended Evolutionary Synthesis: where does Evo-Devo fit in? En: *Theory in Biosciences* 137:169-184.
- [2] J. Perona, Á., (2015). Criticism, Philosophy and Normativity: Observations for a Controversy, *Revista Portuguesa de Filosofia*, 2015, 188-210.
- [3] — (2016). Pluralism and Soft Rationality in the Philosophy of Ortega and Wittgenstein. En *Ortega y Gasset and Wittgenstein on Knowledge, Belief, and Practice*, Astrid Wagner / Jose Maria Ariso, Berlin-Boston: Rationality Reconsidered, 75-87.
- [4] Kuhn, T., (1970). *The Structure of Scientific Revolutions*. University of Chicago Press.
- [5] (1977). *The Essential Tension*. University of Chicago Press.

The philosophical significance of the molecular interpretation of evolution

Davide Vecchi*
Universidade de Lisboa

There is no doubt that the biochemical and molecular basis of physiological and developmental processes such as DNA replication, protein biosynthesis and morphogenesis has been progressively refined. Since the late 60s, molecular research has also significantly contributed to our understanding of evolutionary dynamics. The tremendous achievement of the molecular understanding of life has been propelled by the development of various technologies. In the 60s, comparative biochemistry played a major role in delineating the characteristic features of the emerging molecular interpretation of evolution. In the last twenty years or so, comparative genomics has also heavily contributed to such delineation. Molecular evolution research today encompasses a large body of experimental and theoretical contributions. Nevertheless, despite some notable exceptions, this extensive body of experimental and theoretical work has not received, in my opinion, the philosophical attention it deserves. Apart from important philosophical contributions dealing with issues such as the nature of variation (Beattie 2010, Huneman 2017, Nuño de la Rosa & Villegas 2019) and of lineage dynamics (Dietrich 1996, Dietrich & Millstein 2008, Razeto-Barry & Frick 2011, Sarkar 2015, Brunet & Doolittle 2018), what is lacking is a systematic philosophical analysis, in particular of the experimental and theoretical contributions produced by evolutionary geneticists and molecular evolutionists in the last 20 years (e.g., Stoltzfus 1999, Stoltzfus 2006, Lynch 2007a, Lynch 2007b, Koonin 2009). In order to fill this conspicuous gap, in this contribution I aim to characterise two general ways in which the molecular interpretation is philosophically significant.

First, the molecular interpretation allows a potentially mechanistic conceptualisation of evolutionary dynamics. King & Jukes' (1969) comparative analysis of amino acid sequences paved the way for a conceptualisation of protein evolution in terms of single mutational events (e.g., amino acid replacements) followed by fixation. Protein evolution in this sense provides an evolutionary "toy model" (Reutlinger et al. 2018) with various epistemological benefits. One benefit is the possibility of decomposing evolutionary causal chains in terms of mutational events. I suggest that mutational events should be characterised in terms of genomic changes of various nature generated by a variety of molecular mechanisms (Koonin 2009) that can be conceptualised in terms of acquisition and loss of "parts". Even though such mutational events might initially have negligible fitness

* davide.s.vecchi@gmail.com

effects (a significant portion of single nucleotide substitutions is synonymous while most of the rest engender “conservative” amino acid substitutions, i.e., not changing protein structure and function in the “normal” developmental context, King & Jukes 1969), they are much more common than originally thought by population geneticists (Kimura 1968) and, in the absence of purifying selection, they might be accumulated during evolution (Lynch 2007a, Koonin 2009) as “excess capacity” (Stoltzfus 1999). A related advantage of the molecular approach lies in clarifying the relationship between development and evolution. The origin of a trait is not identifiable with mutational events but, rather, with their phenotypic manifestation, which leads, potentially, to phenotypic novelty and organismal complexity. Stoltzfus (2006) has proposed to conceptualise the process of phenotypic manifestation in terms of “mutation-and-altered-development”: it is only within a specific organismal system embedded in a specific environmental context that the developmental potential embodied by accumulated “excess capacity” can be actualised by producing sortable phenotypes. These phenotypes might be molecular (e.g., new proteins, new genetic networks) or not. More generally, they might be conceptualised in terms of acquisition or loss (McShea & Hordijk 2013) of organismal “parts”. Stoltzfus’ analysis is mechanistic in ethos, as understanding the process of “mutation-and-altered-development” requires recomposition and situating analyses (Bechtel & Abrahamsen 2009). So conceived, the origin of novel traits makes development central to evolution, as all developmentalist approaches to evolution, in my opinion correctly, stress (Mahner and Bunge 1997).

The other major philosophical benefit of molecular approaches concerns the possibility of rigorously testing adaptive hypotheses. Two challenges stemming from molecular studies are pivotal in this sense. Since the 60s, the mismatch between the generally higher and constant rates of molecular evolution and the generally slower and inconstant rates of phenotypic evolution (the “molecular-morphological paradox”, cf. Dietrich 1998) was recognised (Suárez-Díaz 2009). Recent molecular studies not only reinforce the significance of this challenge to adaptationism (Lynch 2007a, Koonin 2009), but add an important twist by identifying a significant mismatch between unicellular and multicellular lineages’ evolutionary dynamics: in the case of multicellular organisms, low effective population sizes prevent purifying selection from curbing the accumulation of slightly deleterious genomic changes. The ensuing “genomic challenge to adaptationism” (Sarkar 2015) re-proposes the crucial question of what is the role of selection in the evolution of the seemingly adaptive phenotypes displayed by multicellular organisms, those that since Darwin have been the central focus of evolutionary theory. Within this context, some molecular evolutionists have proposed constructive neutral models (Stoltzfus 1999, Lynch 2007c, Gray et al. 2010). Probably most intriguingly, molecular approaches seem to paint an idiosyncratic picture of evolutionary dynamics that is at odds with the unifying pretensions of the Modern Synthesis and, more generally, the ideal of unification pervading evolutionary biology (Stoltzfus 2017).

Bibliography

- [1] Beattie, J. (2010). Reconsidering the Importance of Chance Variation. In Pigliucci M., Müller, G. (eds.). *Evolution: The Extended Synthesis*. Cambridge: MIT Press.
- [2] Bechtel, W., Abrahamsen, A. (2009). Decomposing, recomposing, and situating circadian mechanisms: Three tasks in developing mechanistic explanations. In H. Leitgeb, & A. Hieke (Eds.). *Reduction and elimination in philosophy of mind and philosophy of neuroscience* (pp. 173-186). Frankfurt: Ontos Verlag.
- [3] Brunet, T.D.P., Doolittle, W.F. (2018). The generality of Constructive Neutral Evolution. *Biology and Philosophy* 33:2 <https://doi.org/10.1007/s10539-018-9614-6>
- [4] Dietrich, M.R. (1996). Monte Carlo Experiments and the Defense of Diffusion Models in Molecular Population Genetics. *Biology & Philosophy* 11:339-356.
- [5] Dietrich, M.R. (1998). Paradox and persuasion: negotiating the place of molecular evolution within evolutionary biology. *Journal of the History of Biology* 31(1):85-111.
- [6] Dietrich, M.R., Millstein, R. (2008). The Role of Causal Processes in the Neutral and Nearly Neutral Theories. *Philosophy of Science* 75(5):548-559.
- [7] Gray, M.W. et al. (2010). Irremediable Complexity? *Science* 330: 920-921.
- [8] Huneman, P. (2017). Why Would We Call for a New Evolutionary Synthesis? The Variation Issue and the Explanatory Alternatives. In P. Huneman and D. M. Walsh (eds), *Challenging the Modern Synthesis: Adaptation, Development, and Inheritance*, Oxford: Oxford University Press, pp. 68-110.
- [9] Kimura, M. (1968). Evolutionary Rate at the Molecular Level, *Nature* 217:624-6.
- [10] King, J. L., Jukes, T. H. (1969). Non-Darwinian Evolution: Most Evolutionary Change in Proteins May Be due to Neutral Mutations and Genetic Drift. *Science*, 164: 88-98.
- [11] Koonin, E. V. (2009). Darwinian evolution in the light of genomics. *Nucleic Acids Research*, 37(4):1011-1034.
- [12] Lynch, M. (2007a). The frailty of adaptive hypotheses for the origin of organismal complexity. *PNAS* 104: 8597-8604.
- [13] Lynch, M. (2007b). *The origins of genome architecture*. Sunderland: Sinauer.
- [14] Lynch, M. (2007c). The Evolution of Genetic Networks by Non-adaptive Processes', *Nature Reviews Genetics* 8:803-13.
- [15] Mahner, M., and Bunge, M. A. (1997). *Foundations of Biophilosophy*. Berlin: Springer.

- [16] McShea, D. W., Hordijk, W. (2013). Complexity by subtraction. *Evolutionary Biology*, 40(4): 504-520.
- [17] Nuño de la Rosa, L. & Villegas, C. (2019). Chances and Propensities in Evo-Devo. BJPS axz048, <https://doi.org/10.1093/bjps/axz048>
- [18] Razeto-Barry P, Frick R. (2011). Probabilistic causation and the explanatory role of natural selection. *Stud Hist Phil Biol Biomed Sci.* 42(3):344-55. doi:10.1016/j.shpsc.2011.03.001.
- [19] Reutlinger, A. et al. (2018). Understanding (With) Toy Models. *British Journal for the Philosophy of Science* 69(4):1069-1099.
- [20] Sarkar S. (2015). The genomic challenge to adaptationism. *British Journal for the Philosophy of Science* 66(3):505-536.
- [21] Stoltzfus, A. (1999). On the possibility of constructive neutral evolution. *Journal of Molecular Evolution* 49:169-181.
- [22] Stoltzfus A. (2006). Mutationism and the Dual Causation of Evolutionary Change. *Evolution and Development* 8:304-317.
- [23] Stoltzfus, A. (2017). Why we don't want another Synthesis. *Biology Direct* 12, 23 doi:10.1186/s13062-017-0194-1
- [24] Suárez-Díaz, E. (2009). Molecular evolution: concepts and the origin of disciplines. *Studies in History and Philosophy of Biological and Biomedical Sciences* 40:43-53.

FILOSOFÍA Y METODOLOGÍA
DE LA CIENCIA

The arrival of the Smartest

Giorgio Airoidi*
UNED

The turn of the millennium has brought novel and fecund theories in the areas of both evolutionary biology and of philosophy of mind. In the former, the preponderant role of selection in evolution has been challenged, and alternative mechanisms have been proposed as sources of novelties. In the latter, the representational theories of mind have been criticised for being unable to explain all aspects of human cognition, that is often inter-dependent on active external resources. The extensions of classical theories in both areas better capture the complex and interwoven realities of life and mind. Nevertheless, novel theories of mind often justify the appearance and fixation of cognitive abilities in terms of adaptation, and pluralistic accounts of evolution often recur to classical representational theories of mind. In this talk, I advocate for the integration of non-adaptationist and non-cognitivist hypothesis, as I believe such an integration could benefit our comprehension of what cognition is and how cognition evolved.

In the first section, I discuss the novel approaches in theory of mind that go beyond the traditional view of cognition as the processing of representations taking place within the neural connections of a computer-like brain (Adams and Aizawa 2001). This ‘brain-bound’ approach is strongly internalist and functionalist – features that suit very well with adaptationist accounts of its evolution – , as it considers intelligence as the ability to solve problems and the brain as a computer that processes symbols according to syntactic rules: an idea behind the first essays on Artificial Intelligence (Reeke and Edelman 1988). A first group of alternative theories of cognition interpret it as ‘embodied’, or as an ability distributed among brain, body and environment (e.g. Clark 2007, Gallagher 2005). Externalist theories (also known as embedded) grant an active role in cognition to the environment, that can just supply inputs and outputs to cognitive processes (weak externalism: Rupert 2004, Wheeler 2010), or have a strong causal role (scaffolded mind hypothesis). In the extended mind hypothesis, the frontiers between brain, body and physical and cultural environment disappear with regards to cognitive states and processes, as all these elements are inter-linked through systemic relationships (Wilson 2004, Clark and Chalmers 1998, Clark 2011).

In the second part of the talk, I review the criticisms to the adaptationist idea that selection is a necessary and sufficient condition for the appearance of novel phenotypic

* gairoidi1@icloud.com

traits. In particular, I illustrate what I consider the three metaphysical pillars behind this view: the reification of the concept of niche (Fodor & Piattelli-Palmarini 2010), the functionalization of the phenotype (Altenberg 2005), the optimization action of selection (Lehman and Stanley 2011). While selection explains adaptation, the novelties upon which selection acts might not always be a result of selection itself. I thus briefly introduce some of the proposals of non-adaptationist sources of variation, sometimes collected under the name of Extended Evolutionary Synthesis (EES) (Laland et al. 2015). Based on the putative source of evolutionary novelties, I classify them as genetic (Kimura 1983, Wagner 2015), phenotypic (Jablonka and Lamb 2005), developmental (Müller 2008) and systemic accounts (Kauffman 1993).

I finally analyse some of the accounts of evolution of cognition available in the literature. The majority of them still relays on a pure adaptationist approach (Geary 2004, Fernando et al. 2012), although they have difficulties to demonstrate that there were differences in cognitive abilities among our remote ancestors, that these differences were heritable, and that they ensured a fitness advantage among individuals of the species, and not between species. These difficulties mainly rely on the lack of direct fossil evidence of cognition: brain size, tools and the like are used as proxy for cognitive abilities (Roth and Dicke 2005, Lande 1979, Jeffares 2010). Non-selective hypotheses avoid these critical issues because they often offer mechanistic and not historical accounts (Cherniak 2009, Dor & Jablonka 2000, Varela et al. 2016).

I conclude stressing that there seems to be little communication between the new frontiers of evolutionary biology and of philosophy of mind: most evolutionary accounts in the philosophy of mind still follow the adaptationist paradigm, while the cognitivist-representational paradigm still reigns in evolutionary biology. We start thinking that intelligence is not just problem-solving, and that evolution is not just, in Darwin's word, war of nature, famine and death. Merging our investigations about 'who the smartest is' and 'where the smartest arrived from' into a unified approach could greatly improve our understanding of both what cognition is and how it evolved.

Bibliography

- [1] Adams, F., Aizawa, K. (2001). The bounds of cognition. *Philosophical Psychology*, 14 (1): 43-64
- [2] Altenberg (2005). Modularity in Evolution: Some Low-Level Questions, in Callebaut, W., Rasskin-Gutman, eds. (2005). *Modularity. Understanding the Development and Evolution of Natural Complex Systems*. The MIT Press: 99-128.
- [3] Cherniak, C., (2009). Brain wiring optimization and non-genomic nativism. In M. Piattelli-Palmarini, M., Uriagereka, J., Salaburu, P. (eds), *Of Minds and Language*:

- a Dialogue with Noam Chomsky in the Basque Country*, Oxford, Oxford University Press, 2009:108-19
- [4] Clark, A. (2007). Re-inventing ourselves: The plasticity of embodiment, sensing, and mind. *The Journal of Medicine and Philosophy*, 32(3): 263-282.
- [5] Clark, A. (2011). *Supersizing the mind: embodiment, action, and cognitive extension*, Oxford: Oxford University Press, 2011
- [6] Clark, A., Chalmers, D. J. (1998), The Extended Mind, *Analysis* 58 (1): 7-19
- [7] Dor, D., Jablonka, E. (2000). From cultural selection to genetic selection: a framework for the evolution of language. *Selection*, 1:33-55
- [8] Fernando, C., Szathmary, E., Husbands, P. (2012). Selectionist and evolutionary approaches to brain function: a critical appraisal. *Frontiers in Computational Neuroscience*. 6 (24): 24.
- [9] Fodor, J., Piattelli-Palmarini, M. (2010). *What Darwin got wrong*. Profile Books
- [10] Gallagher, S. (2005). *How the body shapes the mind*. Oxford: OUP
- [11] Geary, D. C. (2004). *Origin of Mind: Evolution of Brain, Cognition, and General Intelligence*. American Psychological Association (APA)
- [12] Jablonka, E., Lamb, M. J. (2005). *Evolution in Four Dimensions. Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life*. The MIT Press.
- [13] Jeffares, B. (2010). The co-evolution of tools and minds: cognition and material culture in the hominin lineage. *Phenomenology and the Cognitive Sciences*. 9:503-520.
- [14] Kauffman, S. A., (1993). *The Origins of Order. Self-Organization and Selection in Evolution*. Oxford University Press.
- [15] Kimura, M. (1983). *The neutral theory of molecular evolution*. Cambridge University Press
- [16] Laland K.N., Uller T., Feldman M.W., Sterelny K., Müller G.B., Moczek A., Jablonka E., Odling-Smee J. (2015). The extended evolutionary synthesis: its structure, assumptions and predictions. *Proc. R. Soc. B* 282: 20151019.
- [17] Lande, R. (1979). Quantitative Genetic Analysis of Multivariate Evolution, Applied to Brain: Body Size Allometry. *Evolution*, 33(1)2:402-416.
- [18] Lehman, J., Stanley, K. O. (2011). Abandoning objectives: Evolution through the search for novelty alone. *Evolutionary computation*, 19(2):189-223.

- [19] Müller, G. B. (2008). Evo-Devo as a discipline. In Minelli A., Fusco G. (eds.), *Evolving Pathways: Key Themes in Evolutionary Developmental Biology*. Cambridge University Press:5-30.
- [20] Reeke, G. N. Jr., Edelman, G. M. (1988). Artificial Intelligence. *Daedalus*, 117(1): 143-173.
- [21] Roth G, Dicke U. (2005). Evolution of the brain and intelligence. *Trends Cogn Sci.*, 9(5): 250-257.
- [22] Rupert, R. D. (2004), Challenges to the hypothesis of extended cognition, *Journal of Philosophy* 101 (8): 389-428.
- [23] Varela, F. J., Thompson, E., Rosch, E. (2016), *The embodied mind : cognitive science and human experience*. Cambridge, MA : MIT Press. ISBN 9780262529365
- [24] Wagner, A. (2015). *Arrival of the Fittest*, Oneworld Publications
- [25] Wheeler, M. (2010), In Defense of Extended Functionalism, in *Menary* (2010a): 245-270.
- [26] Wilson, R. A. (2004). *Boundaries of the mind: The individual in the fragile sciences-Cognition*. Cambridge University Press.

Papel de los Valores en las Ciencias de Diseño: de la Axiología de la Investigación a la Ética de la Ciencia en Educación

Ana María Alonso*

Centro de Investigación de Filosofía de la Ciencia y la Tecnología, UDC

I

La investigación realizada sobre la Covid-19 ha puesto de manifiesto, una vez más, el nexo entre los valores internos (cognitivos, metodológicos, etc.) y los valores externos (sociales, económicos, ...) de la actividad científica. Esto propicia considerar también los valores éticos, tanto endógenos como exógenos. La base está en un vínculo entre la Metodología de la Ciencia y la Axiología de la Investigación, que paralelamente resalta la relevancia de la Ética de la Ciencia para la Filosofía de la Ciencia.

Tienen los valores especial interés en las Ciencias Aplicadas, en general, y en las Ciencias de Diseño, en particular, debido sobre todo a las consecuencias. El foco del análisis se centra en la Educación por varias razones: (i) su condición de Ciencia de Diseño; (ii) la estrecha imbricación que mantiene con la Tecnología; (iii) una marcada dimensión ética; (iv) su importancia estratégica en la Sociedad del Conocimiento; y (v) por tratarse de una disciplina que apenas ha recibido la atención de la Filosofía de la Ciencia.

II

Al aceptar que la Ciencia está cargada de valores, se consideran también los valores éticos. Están presentes en Educación, que es una disciplina dual, en cuanto Ciencia Social y Ciencia de lo Artificial. El estudio atañe a su dimensión artificial, como Ciencia Aplicada de Diseño que, además, está vinculada a procesos tecnológicos. Si la Ciencia y la Tecnología no son éticamente neutrales, el investigador tiene una responsabilidad ética respecto de su trabajo (Schrader-Frechette, 2005, pp. 51-79).

En cuanto que la Ciencia es una actividad humana libre, hay pautas éticas en la Ciencia Básica (González, 1999). Pero en las Ciencias de Diseño, por ser Ciencias Aplicadas, hay más elementos en liza: a) buscan conocimiento orientado a la solución de problemas concretos; b) sus procesos dependen de los fines buscados; y c) su contenido se evalúa por su relevancia epistémica y, además, por la utilidad práctica. Las Ciencias Aplicadas

* a.alonsor@udc.es

están cargadas de valores de modo diferente a las Ciencias Básicas (Niiniluoto, 2014). Por eso, junto a la Axiología de la Investigación, cobra una mayor dimensión la Ética de la Ciencia, puesto que considera la actividad libre de los agentes: la elección de los fines de la investigación y los medios para llevarla a cabo. Esto se aprecia en la Ciencia de la Educación.

III

El conocimiento científico en Educación está orientado a la práctica educativa, para resolver problemas concretos. La fundamentación filosófica de esta Ciencia guarda relación con valores, como son los principios de carácter ético que orientan esa práctica. Sus procesos involucran valores internos y externos, pero también valores éticos endógenos y exógenos.

Como Ciencia de lo Artificial, la Ciencia de la Educación busca aumentar las posibilidades humanas. Los investigadores de la Educación no se limitan a hacer predicciones acerca de los fenómenos educativos, sino que dan pautas para orientar la actividad educativa para mejorarla. Esas pautas están condicionadas por un conjunto de factores cognitivos (representaciones, valoraciones), volitivos (intereses, deseos y preferencias) y de tipo contextual.

La actividad educativa comporta un desarrollo en el tiempo de acciones humanas intencionales, que está orientada hacia algo éticamente deseable (fines moralmente valiosos). Así, junto a los procedimientos de partida y los métodos de investigación rigurosos (González, 2015, pp. 255-272), hay que considerar lo valioso en términos educativos, pues los fines han de ser relevantes y, de ningún modo, desaconsejables.

Hay pues una escala de valores en esta actividad intencional, desarrollada por agentes intencionales. Esos valores propician unos objetivos, que deben ser viables en un contexto social. Desde la perspectiva interna, la viabilidad depende de tener los medios adecuados para realizar los fines propuestos; y, desde la vertiente externa, hay que considerar los resultados esperados y las consiguientes consecuencias. Los fines han de ser razonables, según valores relacionados con las soluciones propuestas y su repercusión en el entorno social.

IV

Debe distinguirse en la Ciencia, entre la actividad humana y el contenido cognitivo. En las Ciencias de Diseño, para alcanzar nuevos retos, se requiere sintetizar conocimiento. Las teorías y modelos –al igual que en las Ciencias Básicas– se juzgan primero respecto de sus utilidades epistémicas. Pero en la Ciencia Aplicada hay también utilidades prácticas. Así, en la Educación como Ciencia Aplicada, el conocimiento sintetizado está

orientado a resolver problemas de la práctica educativa. Hay una dimensión cognitiva y una significación práctica en la propia síntesis artificial.

Es en la vertiente práctica donde inciden especialmente los valores, entre ellos los éticos: (i) para determinar el éxito práctico del conocimiento; y (ii) la valoración de las consecuencias de los resultados. Pero también (iii) como recomendaciones orientadoras de las acciones, criterios con los que puede comprometerse el investigador (a veces son objeto de disputas éticas); y (iv) criterios asociados a las condiciones y circunstancias desde las que se juzga la propia acción. Así, el científico es moralmente responsable de los resultados de esta actividad.

Junto a la tarea metodológica de la Educación, hay que considerar entonces los valores: (a) al seleccionar los objetivos de los diseños, cuando se analizan los problemas planteados; (b) cuando se proponen modelos científicos para integrar hipótesis, dentro del entorno de teorías y bajo el influjo de marcos teóricos más amplios; y (c) cuando se contrastan las propuestas con la experiencia educativa.

Mediante la investigación educativa cabe: (i) encontrar soluciones a problemas que no son puramente cognitivos; (ii) dar con modelos científicos que no se reducen a acoplar hipótesis dentro de un entorno teórico más amplio; y (iii) alcanzar contrastaciones que acompañan a las personas y la sociedad, donde la observación y la experimentación deben repensarse desde la aceptación de valores. Esto involucra a la Axiología de la Investigación y también a la Ética de la Ciencia.

Por una parte, la Ciencia de la Educación aporta conocimiento acerca de su objeto de estudio –la actividad educativa–, pero su desarrollo metodológico, en su perspectiva interna, está entrelazado con una pluralidad de valores. Por otra parte, desde la perspectiva externa, el entorno condiciona las conclusiones de las investigaciones educativas. Para adaptarse al entorno, el diseño educativo (con sus objetivos, procesos y resultados) debe contemplar aspectos sociales, culturales, políticos, etc. A esto se unen otros aspectos que funcionan como condición de posibilidad, como la legislación vigente en cada momento y el entorno tecnológico.

V

La relación de las Ciencias de Diseño con los valores, tanto internos como externos, se reafirma como una cuestión filosófica central. Así, la moralidad no puede estar desconectada de la Ciencia (Rescher, 1999, p. 67). Esto es especialmente relevante en la investigación educativa, por incidir en la componente metodológica, pero también por su vínculo creciente con la Tecnología.

Para resolver problemas concretos, la Educación depende de la creatividad científica y, cada vez más, de las innovaciones tecnológicas. La Tecnología (TICs), que fue inicialmente “externa” a la Ciencia de la Educación, ahora es una condición de viabilidad.

Los diseños hacen posibles nuevos objetivos, procesos y resultados; para ser viables y eficientes, necesitan desarrollos tecnológicos (González, 2021). Esto involucra nuevas cuestiones éticas respecto de la actividad tecnológica en sí misma considerada y el modo como amplía el campo educativo con nuevas posibilidades, donde hay consecuencias con un impacto social incierto.

Mediante el nexo entre creatividad e innovación se potencia la componente artificial de una actividad humana dirigida a fines, que trabaja a base de diseños y tiene una clara dimensión social. En este contexto el papel de los valores se convierte en eje del perfeccionamiento humano en un medio social y cultural. Se transita así de los valores de la Educación a la educación en valores.

Referencias

- [1] González, W. J. (1999). Ciencia y valores éticos: De la posibilidad de la Ética de la Ciencia al problema de la valoración ética de la Ciencia Básica. *Arbor*, v. 162 (638), 139-171.
- [2] González, W. J. (2015). *Philosophico-Methodological Analysis of Prediction and its Role in Economics*. Dordrecht: Springer.
- [3] González, W. J. (2021). A Framework for the Ethical Analysis of the Internet and Economics: Three Focuses, the Role of Economics and the Ethical Values as a Science”. En Rodríguez Valls, F. y Padial, J. J. (eds), *Ciencia y Filosofía: Estudios en Homenaje a Juan Arana*. Sevilla: Thémata. En prensa.
- [4] Niiniluoto, I. (2014). Values in Design Sciences. *Studies in History and Philosophy of Science*, 46, 11-15. <http://dx.doi.org/10.1016/j.shpsa.2013.11.002>
- [5] Rescher, N. (1999). *Razón y valores en la Era científico-tecnológica*. Barcelona: Paidós.
- [6] Shrader-Frechette, K. (2005). Objectivity and Professional Duties Regarding Science and Technology. En González, W. J. (ed.), *Science, Technology and Society: A Philosophical Perspective* (51-79). A Coruña: Netbiblo.

Vida como género natural: dificultades metafísicas en la definición de vida

Alba Amilburu Martínez*
Universidad Pública de Navarra

Jaime Soler Parra**
Universidad Nacional de Educación a Distancia

1 Introducción

En la actualidad existen un gran número de propuestas para definir vida (Popa, 2004). Esta cuestión ha cobrado gran relevancia de la mano de nuevas disciplinas como la astrobiología, la biología sintética o la vida artificial. Ahora bien, la proliferación de propuestas no ha supuesto un mayor consenso a la hora de definir el concepto, más bien al contrario.

No es de extrañar, por tanto, que diversos autores se hayan mostrado escépticos respecto a la posibilidad de definir vida (Tirard et al., 2010; Cleland, 2012, Machery 2012). Estos autores consideran que definir vida es encontrar qué propiedad esencial, o propiedades esenciales, permiten caracterizar de forma necesaria y suficiente vida. Para todos ellos vida se define como género natural esencialista, con un fuerte compromiso metafísico.

Es cierto que no todos estos autores consideran estos compromisos desde una misma perspectiva. Así, Machery (2012) insiste en que vida no puede ser entendida como concepto científico, en la medida que no existe ninguna propiedad asumida por todas las disciplinas que definen vida. Por su parte, Tirard et al. (2010) señalan que vida es un concepto histórico, y como tal rehuye de ser definido. Más que definiciones, estos autores señalan la relevancia de los contextos de investigación.

Probablemente las críticas más contundentes sean las de Cleland (2012). Para esta autora definir vida supone confundir el plano lingüístico con el epistémico. Además, considera imposible definir el concepto de vida en un contexto como el actual, donde sólo se conoce un tipo de vida, esto es, vida en la Tierra. En definitiva, todos estos autores apuntan a que definir vida es considerarla como un género natural esencialista. Esto supone asumir unos fuertes compromisos metafísicos. Ahora bien, cabe preguntarse ¿es este modo fuertemente esencialista el empleado por los investigadores para definir vida?

* alba.amilburu@gmail.com

** jaimesolerparra@yahoo.es

2 Alternativas no esencialistas de vida como género natural

Algunos autores, entre los que destacan Bich y Green (2018), han puesto de manifiesto que los científicos emplean las definiciones como herramientas epistémicas y teóricas, útiles para la investigación científica, abiertas y flexibles. A estas definiciones, alejadas de las definiciones comprometidas metafísicamente, es a lo que han venido a llamar como definiciones operacionales. A través del estudio de casos, muestran cómo los investigadores emplean las definiciones como guías de investigación, y pueden modificarlas conforme a los resultados de tales investigaciones. Desde una perspectiva diferente otros autores, como Amilburu et al. (2020), muestran también el valor heurístico de las definiciones de vida.

Por otro lado, debe tenerse en cuenta que, frente al modo fuertemente esencialista de vida como género natural, diversos autores ponen de manifiesto la existencia de modos de entender el concepto de género natural menos comprometidos metafísicamente. Sería el caso de Diéguez (2013), para quien vida podría entenderse como una agrupación homeostática de propiedades, Ferreira y Umérez (2018), con su propuesta de vida como agrupación estable de propiedades, o Soler (2019), para quien vida sería un género natural promiscuo.

Todas estas alternativas se caracterizan por debilitar los compromisos metafísicos de los géneros naturales esencialistas, aunque partiendo de distintos supuestos. Así, en el caso de las agrupaciones homeostáticas de propiedades el realismo de la propuesta vendría dado por los mecanismos. En este caso coexisten los compromisos metafísicos, a través de los mecanismos, con la presencia de elementos convencionales.

Ahora bien, autores como Craver (2009) señalan algunas dificultades importantes que deben ser tenidas en cuenta. Para este autor, resulta problemático justificar la independencia metafísica de las agrupaciones homeostáticas a partir del concepto de mecanismo. Esto es así porque dentro de los mecanismos existen de forma inherente elementos convencionales. Por tanto, resulta contradictorio justificar la independencia metafísica a partir de un concepto que no es metafísicamente independiente. De forma muy similar puede criticarse la propuesta de Ferreira y Umerez (2018), en la medida que el concepto de estabilidad no está suficientemente justificado desde una perspectiva metafísica.

Por su parte, la propuesta de vida como género natural promiscuo prescinde de compromisos metafísicos, centrándose en aspectos epistémicos. Al prescindirse de los compromisos metafísicos, desaparecen los problemas que surgen en su justificación, que es lo que ocurre en propuestas como las agrupaciones homeostáticas de propiedades. Sin embargo, no todos los autores estarían de acuerdo con esta propuesta. En ocasiones se ha señalado que los géneros naturales promiscuos no pueden distinguirse de los géneros convencionales, ya que no queda suficientemente justificado conforme a qué criterio se consideran como realistas.

3 Los compromisos metafísicos y el dualismo realismo/convensionalismo

De los puntos anteriores se pueden deducir algunas cuestiones relevantes. La primera de ellas es que los investigadores emplean las definiciones en un sentido operacional (Bich y Green, 2013), lo cual es compatible con modos no esencialistas de considerar los géneros naturales. Sin embargo, el problema respecto a la justificación del estatus metafísico de vida como género natural permanece (Amilburu y Soler, 2021).

A pesar de la insistencia en la gradualidad de vida, la separación entre los elementos reales y convencionales sigue estableciéndose de forma dicotómica. Lo real queda radicalmente separado de lo convencional a través de una clara línea de demarcación. Pero esto entra en conflicto con el hecho de que los aspectos reales y convencionales aparecen intrínsecamente unidos, al menos en conceptos complejos como el de vida. La cuestión que se mantiene es ¿puede la idea de género natural preservar el estatus ontológico que se le atribuye renunciando a esa carga metafísica, teniendo en cuenta que es aquello que lo justifica?

Es cierto que la propuesta promiscua trata de trascender esa disyuntiva. Ahora bien, la insistencia en el realismo de las diferentes propuestas solo puede hacerse desde dos supuestos. El primero de ellos es asumir compromisos metafísicos a partir de las esencias o mecanismos, algo que, como ya se ha señalado, resulta problemático. El segundo es prescindir de tales compromisos, pero entonces, según numerosos autores, se está traicionando al propio concepto de género natural. Por tanto se propone que, si no es posible escapar a la dicotomía realismo/convensionalismo, la única alternativa posible es abandonar los géneros naturales para conceptos complejos como el de vida.

Bibliografía

- [1] Amilburu, A., Moreno, A., Ruiz-Mirazo K. (2020). “Definitions of life as epistemic tools that reflect and foster the advance of biological knowledge”, *Synthese*, 1-21.
- [2] Amilburu Martínez, A., y J. Soler Parra. (2021). “Géneros Naturales Y definición De Vida: Una crítica a La Perspectiva metafísica De Lo Vivo”. *Daimon Revista Internacional De Filosofía*, pp. 201 -6, doi:10.6018/daimon.374851
- [3] Bich, L., Green S. (2018). “Is Defining Life Pointless? Operational Definitions at the Frontiers of Biology”, *Synthese*, 195(9), pp. 3919-3946.
- [4] Cleland, C. E. (2012; 2011). “Life without definitions”, *Synthese*, 185(1), 125-144.
- [5] Craver, C. F. (2009). “Mechanisms and natural kinds”, *Philosophical Psychology*, 22 (5), pp. 575-594.

- [6] Diéguez, A. (2013). “Life as a homeostatic property cluster”, *Biological Theory*, 7(2), pp. 180-186.
- [7] Dupré, J. (1999). “Are whales fish”, en: D. Medin, & S. Atran (Eds.), *Folkbiology* (pp. 461-476) MIT Press
- [8] Ferreira, M., Umerez, J. (2018). “Dealing with the changeable and blurry edges of living things: a modified version of property-cluster kinds”, *European Journal for Philosophy of Science* 8, pp. 493–518.
- [9] Machery E. (2012). “Why I Stopped Worrying about the Definition of Life... and Why You Should as Well”, *Synthese* 185(1), pp. 145–164
- [10] Popa, R. (2004). *Between Necessity and Probability: Searching for Definition and Origin of Life*. Heidelberg: Springer.
- [11] Tirard, S., Morange, M., Lazcano, A. (2010). “The Definition of Life: A brief History of an Elusive Scientific Endeavor”, *Special Collection of Essays: What is Life? Astrobiology*, 10(10), pp. 1003-1009.

Framing the Epistemic Schism of Statistical Mechanics

Javier Anta*

Universitat de Barcelona, LOGOS, BIAP.

Abstract

In this talk I present the main results from Anta (2021), namely, that the theoretical division between Boltzmannian and Gibbsian statistical mechanics should be understood as a separation in the epistemic capabilities of this physical discipline. In particular, while from the Boltzmannian framework one can generate powerful explanations of thermal processes by appealing to their microdynamics, from the Gibbsian framework one can predict observable values in a computationally effective way. Finally, I argue that this statistical mechanical schism contradicts the Hempelian (1958) thesis that the predictive power of a scientific theory is directly proportional to its explanatory potential, and vice versa.

1 Introduction

Undoubtedly, statistical physics is one of the main domains of modern physics. Particularly, statistical mechanics is devoted to the study of the macroscopic behavior of certain systems, such as a gas expanding in a vessel, from their molecular microdynamics. Unlike other physical domains, classical statistical mechanics not only does not possess a single canonical theoretical formalism, but moreover there is no consensus within the philosophical literature about its fundamental problems. Both the basic concepts and procedures underlying statistical mechanics depend on either adopting a Boltzmannian or a Gibbsian perspective.

On the one hand, in Boltzmannian statistical mechanics (BSM) the particular macrostate Γ_M of the system is defined from a set of macrovariables f_i that describe the system observationally. An individual macrostate x of the system is determined by a point in the system's phase space Γ . Therefore, according to BSM the exact state of a molecular system can mechanically described by a single macrostate Γ_M at each time t . The approach of a system to its state of thermal equilibrium can be represented in BSM by a phase path γ originated in a small macrostate Γ_M and directed towards larger phase regions until eventually reaching the phase region associated with the equilibrium macrostate Γ_{Meq} at the

* antajav@gmail.com

moment when e.g., a gas occupies a whole vessel. On the other hand, Gibbsian statistical mechanics (GSM) is based on a set of virtual copies of the same system called ‘ensemble’ and represented by a probability density ρ . The fact that the Gibbs fine-grained entropy (proportional to the measure of ρ) of the density ρ is invariant goes against the second law of thermodynamics, but can be numerically fixed by introducing a coarse-graining procedure to allow entropy to ‘increase’ in the system’s evolution (Uffink, 2007).

2 The Schism between Boltzmann and Gibbs Statistical Mechanics

One of the most illustrative cases of the theoretical untranslatability between BSM and GSM is what is known in the literature as ‘Sklar’s dilemma’ (Sklar, 1993), which can be described for the case of the equilibrium approach of the systems as follows. While for the Gibbsian approximation, once the ensemble ρ reaches the statistical-ensemble equilibrium state, it will remain there if there is no external intervention, according to thermodynamic predictions. For the Boltzmannian approximation, once the individual system reaches its equilibrium macrostate-region Γ_{Meq} , it may eventually fluctuate outside that equilibrium macrostate, against the thermodynamical regularities. This theoretical split has recently been interpreted by Frigg and Werndl (2019) so that BSM and GSM would constitute different types of statistical mechanical theories. Namely, the first as a fundamental theory and the second as an effective theory. Developing on this idea, I argue (Anta 2021) that BSM is a theory capable not only of providing true descriptions but also of generating powerful explanations of thermal behaviors; while GSM is a statistical mechanical theory capable of generating effective predictions about the evolution of the system’s macrovariables.

3 Explanatory Power and Predictive Efficacy

Following Uffink (2007), any plausible statistical mechanical explanation of non-equilibrium phenomena also requires explaining (or at least justifying consistently) why this particular dynamic is somehow theoretically and experimentally compatible with the initial conditions proposed to describe the values of the system at the beginning of this evolution. Then:

Explanatory Power Criterion:

A statistical mechanical description D of the thermal behavior θ of a molecular system constitutes a powerful explanation of θ if this description is abducted from (i) the system’s initial conditions and (ii) the microdynamical properties P of the system that generates such thermal behavior θ .

Firstly, one can show how any GSM-description of the thermal behavior will not meet the above explanatory power criterion because the GSM descriptions of thermal-macroscopic behavior do not specify any microstatistical properties from which can be inferred the macrovariables f_i encoded in ρ^* . GSM simply specify the theoretical procedures required to adjust these values according to thermodynamics. Thus, standard GSM-descriptions fail to satisfy condition (ii) of the above minimum criteria. Since these same Gibbsian descriptions depend constitutively on eliminating the initial system condition by introducing a new probability distribution ρ^* . Therefore, they would also fail to satisfy condition (i). On the other hand, one can argue that certain BSM descriptions of thermal behavior can meet the above explanatory power criterion. According to BSM canonical proposal (Sklar 1993), the approach to equilibrium could be explained from (i) the fact that every possible microstate x contained in low-entropy macrostate Γ_M transits typically towards the equilibrium macrostate Γ_{Meq} ; and from (ii) the fact that the system is initially within a low-entropy macrostate Γ_M . The explanatory potential of the Boltzmannian theoretical apparatus rests directly on the descriptive exploitation of the microdynamics of ‘individual’ physical systems, explanatorily accounting for thermal behaviors such as the free expansion of a gas. Then, BSM stands as a physical theory capable of generating powerful statistical mechanical explanations of certain thermal behaviors. Now, a second criterion can be defined:

Predictive Effectiveness Criterion:

A description of the future macrovariables f_i of a system with thermal behavior θ constitutes an effective prediction if (i) the computation of such values can be computationally treatable, and in that case (ii) if such description specifies a well-defined algorithmic procedure that allows calculating such values by means of realistic computational resources with respect to actual practices.

Although explanatorily powerful and conceptual consistent, most of Boltzmannian predictions of physically significant n -molecular ($n \approx 10^{23}$) systems do not meet the above effectiveness criterion based on mathematical tractability and pragmatic feasibility. The main reason underlying BSM’s inability to predict is its epistemic dependence on individual microstates as fundamental representational structures. Then, BSM predictions are both computationally difficult to perform and extremely computationally expensive. On the other hand, it could be argued that in the context of GSM, descriptions of future macrovariables of the system would be assumed as effective SM-predictions because of the role of statistical techniques like coarse-graining and phase averaging. Therefore, the predictive potential of the Gibbsian apparatus rely on exploiting statistical techniques, mainly in calculating observable values. It allows GSM to generate pragmatically effective and algorithmically-specified predictions of thermal behavior such as the expansion of a gas in a vessel.

4 Epistemic Schism and Hempel's Symmetry Thesis

Each of these frameworks (BSM or GSM) fulfill only one of the two core epistemic functions in the statistical mechanical context (Anta 2021): either to powerfully explain (BSM) or to effectively predict (GSM) thermal behavior. This epistemic schism between the main approaches to classical statistical mechanics that I have defended throughout this talk constitutes a clear counterexample to what is known in the literature as the 'symmetry thesis' (Douglas, 2009) which states (i) that every satisfactory explanation also has predictive potential regarding the phenomenon explained, and (ii) that every satisfactory prediction will have explanatory potential with respect to the predicted values. This was originally defended by Carl Hempel (1958) in his famous text 'The theoretician's dilemma'.

Here, one face a fundamental dilemma: either the statistical mechanical apparatus is conceptually consistent enough to generate descriptions that powerfully explain thermal phenomena from the microdynamics of their constituents (BSM-horn), or the statistical mechanical apparatus is pragmatic enough to generate descriptions that efficiently predict thermodynamic values (GSM-horn). When a description like D_1 satisfactorily meets the necessary condition to constitute a powerful explanation, then this SM-description D_1 won't be able to satisfactorily meet the necessary condition of the minimalist criterion to constitute an effective prediction, and the other way around. Therefore, the explanatory potential of statistical mechanics does not correlate directly with the predictive potential of that same theory, against Hempel's (1958) symmetry thesis. Nevertheless, the thesis (Anta 2021) that I have defended in this talk can contribute to understanding how the explanatory powerful conceptual architecture of BSM can be fundamental in accounting for the lack of explanatory power of GSM, just as the predictively effective technical apparatus of GSM can illuminate the practical difficulties that BSM entails. This leads us to think that only by delving into the space opened by the schism between these two statistical mechanical programs can we continue to progress in the rich foundations that underlie this fruitful physical theory.

References

- [1] Anta, J. (2021). The Epistemic Schism of Statistical Mechanics. *Theoria* 36(3): 399-419.
- [2] Douglas, H. (2009). Reintroducing prediction to explanation. *Philosophy of Science* 76 (4), 444-463.
- [3] Frigg, R. & Werndl, C. (2019). Statistical Mechanics: A Tale of Two Theories. *The Monist* 102 (4), 424-438.
- [4] Hempel, C. (1958). The theoretician's dilemma: A study in the logic of theory construction. *Minnesota Studies in the Philosophy of Science* 2, 173-226.

- [5] Sklar, L. (1993). *Physics and Chance: Philosophical Issues in the Foundations of Statistical Mechanics*, Cambridge: Cambridge University Press.
- [6] Uffink, J. (2007). Compendium of the foundations of classical statistical physics. In J. Butterfield, & J. Earman, (Eds.), *Handbook for philosophy of physics*. Amsterdam: Elsevier.

¿Es la convexidad una condición necesaria de las propiedades (naturales) cromáticas en la ceguera al color? Un estudio de caso desde la filosofía experimental

Carlos Arca*
Universidad Complutense de Madrid

Abstract: Gärdenfors sostiene en su modelo de los 'espacios conceptuales' que la convexidad es una condición necesaria de las regiones de los espacios conceptuales para ser 'propiedades naturales'. Un aparente contraejemplo a esta hipótesis es el caso del espacio conceptual cromático de las personas con ceguera al color. Este trabajo presenta los resultados de un estudio de caso único como ejercicio y ensayo de filosofía experimental para dilucidar si en ningún caso de ceguera al color es la convexidad una condición necesaria del espacio conceptual cromático y así comprobar si esta determinación perceptiva es homogénea y consistentemente un contraejemplo a la hipótesis de la teoría de Gärdenfors.

El modelo de los '*espacios conceptuales*' ('*conceptual spaces*') de Peter Gärdenfors (2004, 2005, 2019) ha sido una de las propuestas ha tener en cuenta en las dos últimas décadas por distintas disciplinas. En la filosofía su operatividad se dirige hacia cuestiones relativas al conocimiento y a la epistemología, a la semántica (principalmente de los términos de género natural) y a las posibilidades de resolver algunos problemas filosóficos (como la naturaleza de los predicados proyectables o la analiticidad de proposiciones y enunciados). En ciencias cognitivas es puesto en juego para construir el armazón teórico y dirigir las estrategias epistemológicas y metodológicas de la investigación empírica en torno a la conceptualización, así como alternativa al paradigma simbólico (cf., Fodor 1981) y al paradigma conexionista (cf., Rumelhart y McClelland 1986).

Estos espacios conceptuales son una elaboración desarrollada a partir de los '*quality spaces*' de Quine (2001: 95-97, 2002: 157-158) —probablemente la única concesión de este al innatismo de corte chomskiano (a su vez deudor de la noción de 'innato' presente en la biología de Konrad Lorenz [1980]). La propuesta de Gärdenfors plantea la posibilidad de modelar los conceptos en base a dimensiones cualitativas (D_i) que definen los espacios conceptuales como clases de dimensiones D_1, \dots, D_n . Para ello introduce un utillaje matemático que permite definir métricas y funciones de distancia entre los objetos (puntos en el espacio conceptual, representados como vectores $v = \langle d_1, \dots, d_n \rangle$ con

* carlosarcadelgado12@gmail.com

valores para cada dimensión) lógicamente posibles que caen bajo los conceptos. Y así establecer una noción geométrica de similaridad.

Uno de los ejemplos presentados por Gärdenfors de conceptos modelables de tal forma, y que es usado por este para ilustrar el propio modelo, es el de conceptos de color. Para el autor, estos, en tanto que ejemplos paradigmáticos de conceptos de género natural, se caracterizan por constituir regiones convexas del espacio conceptual de los colores. Regiones convexas que define como regiones en las que todos los casos o instancias de color que se encuentren ‘entre’ dos instancias que caen bajo un mismo concepto de color, caerán igualmente bajo este concepto de color. Formalmente: un conjunto C de puntos (objetos lógicamente posibles) en un espacio S es convexo si y solo si, para todos los puntos x e y en C , cada punto entre x e y está también en C . En términos de una función de distancia $d(x, y)$: dada esta en S , una región C en S es convexa si y solo si $\forall x, y \in C: d(x, y) \in C$. Se define así una ‘propiedad natural’ como una región convexa de un espacio conceptual (Criterio P).

El espacio conceptual de los colores estaría formado así por las dimensiones (a) tono, (b) brillo y saturación, que son isomórficas, respectivamente:

- (a) Del círculo o anillo cromático (popularizado por Goethe). Por tanto, es una dimensión continua y no discreta, y cerrada (circular o circunferencial).
- (b) De un segmento de recta. Por lo tanto, es una dimensión lineal con dos puntos límite o extremos (correspondientes al blanco y al negro).
- (c) De un intervalo de la recta real.

Una de las críticas que se le puede hacer a este planteamiento (y que me consta que se le hace) es el caso de la percepción de los colores, su conceptualización y su clasificación por los individuos que presentan afecciones de ceguera al color (discomatopsia) en alguna de sus variedades (protanopia, deuteranopia, tritanopia y acromatopsia).

La ceguera a alguno de los colores puede suponer un contraejemplo empírico a la convexidad de las regiones y del espacio conceptual de los colores (‘hipótesis relativista’). En principio, la ausencia de la percepción de al menos un color parece segmentar la continuidad entre colores en su consideración fenoménica (al menos en la dimensión del tono) y de ello se podría seguir la no-convexidad de las propiedades cromáticas en estos casos. Sin ser esto óbice para dejar de considerar los colores una propiedad natural bajo tales circunstancias. Pero desde la teoría de los espacios conceptuales se podría defender que el resultado o efecto de las discomatopsias sería más bien un espacio conceptual cromático diferente al que regularmente constituye la conceptualización, al darse una circunstancia y determinación no habitual en la percepción, pero con regiones igualmente convexas (aunque quizás no conectadas) al menos en la percepción. A falta de una relación

lógica entre la ceguera al color y la refutación o ratificación de la convexidad de las propiedades cromáticas, una alternativa es la aproximación experimental. El doble objetivo inicial de la comunicación es presentar un ejercicio de filosofía experimental que, por un lado, sirva de abono al cultivo de la creatividad experimental en filosofía y, por otro lado, exponga el resultado de tratar de falsar la convexidad de las regiones del espacio conceptual de los colores para el caso de ciegos al color. Siendo así que, en el caso de refutarse, se impugne automáticamente la convexidad para este espacio conceptual como una condición necesaria (sin excepciones) de las propiedades naturales. Todo ello pone de manifiesto y supone que, como Gärdenfors (2019) sostiene, la convexidad es una ley empírica y, por tanto, contrastable para cada caso.

Al tener la investigación por objetivo último, bien refutar la universalidad de esta condición necesaria, bien falsar que la propia naturaleza de la ceguera al color deshecha esta universalidad, el trabajo empírico ha consistido en un estudio de caso único con un individuo afectado por un tipo de discomatopsia. Pues tanto si en el resultado se comprueba la convexidad mencionada como si no, se sigue, respectivamente, (i) que el espacio conceptual para ciegos al color es al menos en un caso acorde con la condición de convexidad y por lo tanto la afección no es por naturaleza garante de la invalidez del modelo en todos los casos de discomatopsia, o (ii) que habiendo al menos un caso discordante con la hipótesis de la convexidad como condición necesaria sin excepciones esta queda inmediatamente falsada.

La metodología seguida ha consistido en solicitar al individuo tres tareas:

1. Ordenar una serie de colores variantes en tono (presentados sucesivamente y de forma aleatoria) en base a la similaridad que percibe entre ellos. Ello sin requerirle la enunciación de términos de color, con la intención de obtener información sobre la percepción y no sobre la categorización (lingüística).
2. Clasificar, primero con cuatro y después con ocho términos, las mismas muestras de colores presentadas sucesiva y aleatoriamente.
3. Clasificar, primero con cuatro y después con ocho términos, las mismas muestras de colores presentadas todas de una vez en el orden resultante de la tarea 1.

Las tareas se han llevado a cabo en días distintos, distanciados por al menos un mes. De tal modo que se reduzca en la mayor medida de lo posible el efecto de la realización de una tarea sobre la realización de las demás, al mermar cualquier continuidad o persistencia que favorezca a una mínima potenciación a largo plazo o a la regla de Hebb sobre la neurobiología de la memoria a largo plazo (tanto procedimental como declarativa).

Trabajos citados

- [1] Fodor, J. A. (1981). *Representations*. Cambridge, MA: MIT Press.
- [2] Gärdenfors, P. (2004). *Conceptual Spaces: The Geometry of Thought*. Cambridge: MIT Press.
- [3] Gärdenfors, P. (2005). *The Dynamics of Thought*. Dordrecht: Springer.
- [4] Gärdenfors, P. (2019). “Convexity is an Empirical Law in the Theory of Conceptual Spaces: Reply to Hernández-Conde”. En M. Kaipainen, F. Zenker, A. Hautamäki, y P. Gärdenfors, *Conceptual Spaces: Elaborations and Applications* (págs. 77-80). Cham: Springer.
- [5] Lorenz, K. (1980). *Evolución y modificación de la conducta*. (C. Gerhart, Trad.) México DF: Siglo XXI.
- [6] Quine, W. (2001). *Palabra y objeto*. Barcelona: Labor.
- [7] Quine, W. (2002). “Géneros naturales”. In W. V. Quine, *La relatividad ontológica y otros ensayos* (M. Garrido, y J. L. Blasco, Trans., pp. 147-176). Madrid: Tecnos.
- [8] Rumelhart, D. E. y McClelland, J. (1986). *Parallel Distributed Processing: Exploration in the Microstructure of Cognition*. Cambridge: MIT Press.

An agenda for the philosophy of biological computation

Lorenzo Baravalle*
Faculdade de Ciências da Universidade de Lisboa

Computational techniques and methods are ubiquitous in biology. One may observe that this situation is common to most contemporary scientific fields. Still, the relation between biology and computation is somehow tighter than the relation between computation and other sciences (with the possible exception of cognitive sciences and AI). As a matter of fact, it is not just the case that virtually all biological systems and processes (genetic, developmental, evolutionary, ecological etc.) can be understood, analysed and simulated computationally. Also, the other way round, computational techniques are frequently inspired in biological knowledge (i.e., genetic algorithms, evolutionary computation, etc.). Far from being working tools just for computer scientists, many biologically-inspired algorithms have been employed, in their turn, to model and understand biological phenomena. Finally, and perhaps most notably, computation in biology is not only conceived as a mean to predict the behaviour of a system: biological systems themselves are often interpreted as capable of performing computations in their own right. In this sense, computation provides (similarly, again, to what happens in the disciplines which study mind and cognition) a fundamental tool for theoretical projects aimed to reproduce and control living processes (like Artificial Life or synthetic biology).

My goal in this paper is to offer a broad classification of different approaches to computation in biology and highlight some conceptual issues related to each of them. Some of these issues have already been calling the attention of biologists and philosophers for a long time. Others have passed somehow unnoticed. At any rate, the discussion over these issues have been customarily held within the boundaries of specific philosophical fields (like the philosophy of biology, the general philosophy of science or the philosophy of computation). What is still missing is a broader and systematic account of the relation between biology and computation. I believe that there is a potentially rewarding pay-off, in terms of the intelligibility of certain conceptual problems, in any attempt to rethink traditional debates within new frameworks. It is in this spirit – that is, with the hope of helping to develop further debates in a more focused way – that I try to outline a sort of common agenda for what we may call a philosophy of biological computation (or computing).

* lorenzo_baravalle@yahoo.it

In order to attain this goal, I shall proceed as follow. First, I shall introduce two general distinctions between different uses of computation in science. The first distinction concerns the choice of employing computation as a mean to model phenomena or, alternatively, to explain them. In some cases, computational methods are just tools to solve problems that could not be solved analytically with more traditional mathematical techniques. However, occasionally, scientists conceive a natural system as able to perform computations or behave algorithmically and, accordingly, they try to explain it as a computational system. The second distinction is about the aspects of computational science that are considered as pertinent to, respectively, model or explain natural phenomena. On the one hand, we have research programmes that rely on a computational approach because computer sciences provide them with algorithmic resources and specific mathematical tools. On the other hand, there are areas of natural sciences that interact with computer sciences because these are, at least in part, engineering sciences. In this sense, they offer an important source of ideas to think and build system architectures.

Neither distinction should be considered as completely neat since, in practice, many current theoretical projects adopt hybrids positions. Still, I think that they are helpful insofar as they allow to provide a classification of different approaches to computation in biology. From the four possible combinations between the two distinctions, we obtain four broad categories: mathematical modelling, mathematical explanations, engineering explanations, and engineering modelling. Each category denotes a more or less definite class of research programmes in biology, aimed to stress specific aspects of the interaction between biology and computation. Thus, mathematical modelling includes those theoretical projects (like bioinformatics) that employ computational concepts instrumentally, to build reliable models of biological processes and to test hypotheses statistically. The category “mathematical explanations” includes those proposals, within Artificial Life and system biology, aimed to explain biological processes as abstract computational processes. The approaches which intend to provide engineering explanations of biological systems (like many research programmes within synthetic biology) also seek a computational understanding of biological systems but, instead of focusing on the algorithmic features of living processes, insist on the architectural similarities between organisms and computers. Finally, the label “engineering modelling” denotes those approaches to biology (such as bio-robotics and parts of synthetic biology) which aim to manipulate and modify biological systems and processes through computer engineering.

To conclude my talk, I shall present a short list of conceptual challenges related to the theoretical projects previously presented. More specifically, I shall stress the importance of clarifying the relation between abstract and concrete properties in computational systems, in order to make sense of the multiple realisability of biological properties on non-biological supports. As I shall argue, both mathematical and engineering explanations

vitaly depend on the multiple realisability of biological properties. Without clear criteria about when a biological property is realised, many claims of the supporters of Artificial Life or synthetic biology remains metaphorical. Secondly, I shall emphasise the importance of clarifying some general ontological commitments in the debate over what we can reasonably infer, from the results obtained through the employment of computational techniques, about the nature of biological processes and systems.

What is the folk concept of life?

Claus Beisbart*
University of Bern

Kevin Reuter**
University of Zurich

This paper details the content and structure of the folk concept of life and discusses its relevance for scientific definitions of life. In two empirical studies, we investigate which features of life are considered salient, on the one hand, and universal, on the other. Functionings, such as nutrition and reproduction, but not material composition, turn out to be salient features commonly associated with living beings (Study 1). By contrast, being made of cells is considered a universal feature of living species (Study 2), and, hence, our best candidate for being necessary for life. These results are best explained by the hypothesis that people take life to be a natural kind subject to scientific scrutiny.

Background

What is life? That is, what is characteristic of living beings such as cats, trees, and mushrooms in contradistinction to rocks and robots? This question is by no means new to philosophers, but in the past few decades, it has received increased attention from scientists and ordinary people alike due to the discovery of planets beyond our solar system. Whatever the aim of a particular definition of life, the folk concept of life is crucial for answering the question of what life is. Accordingly, we find plenty of references to the folk understanding of life in the literature. But in the philosophical literature so far, most claims about the concept of life are simply based upon the authors' intuitions and not supported by any empirical evidence. The aim of this talk is to present empirical data on the folk concept of life and to argue that our data suggest that the folk concept of life is a natural kind concept similar to the concept *water*.

* claus.beisbart@philo.unibe.ch

** kevin.reuter@philo.unibe.ch

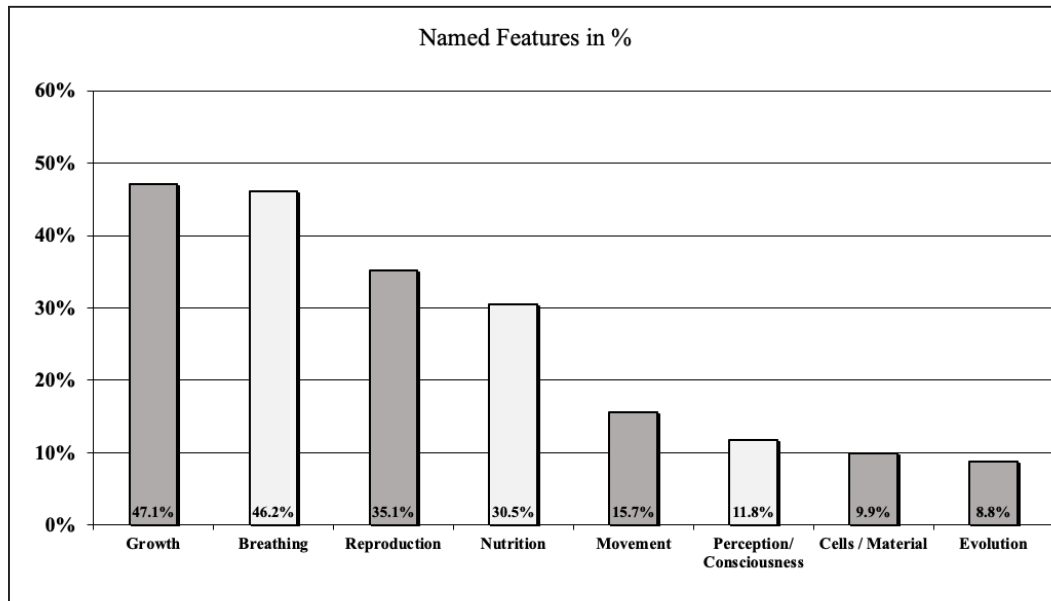


Figure 1: Most frequently named features of life. Frequency is measured relative to people. Thus, 47.1% of people named growth as a feature.

Two Empirical Studies

In Study 1, we used a semantic feature production task to examine which features of life are considered salient by ordinary people. In Study 1, 102 participants were recruited on Prolific and randomly assigned to two different versions of a semantic feature production task. The two versions were given as follows.

- First version: “Which features are characteristic of species of living beings? You can name up to three features”.
- Second version: “Which features do you think distinguish species of living beings from non-living entities?”.

The results depicted in Figure 1 show the fraction of participants who responded with features belonging to the eight categories averaged over both versions of the semantic feature production task. The responses were most frequently classified into the categories growth, breathing, reproduction and nutrition. These are all things that living beings do at the level of a whole living being.

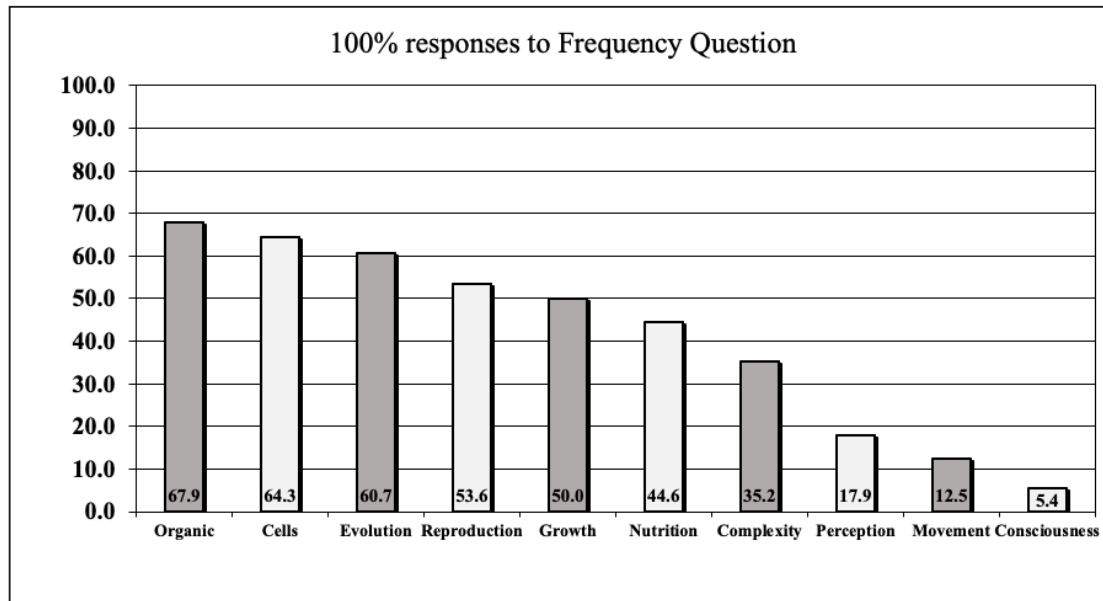


Figure 2: Percentage of participants who think a certain feature is universal, i.e., who responded with 100% in regards to a given feature’s frequency.

In Study 2 we investigated whether there are any features that people consider universal features of life. If a feature is considered universal, then that feature is understood to hold for all species of life. We recruited 127 participants through Prolific. All participants rated the perceived frequency of 10 different features (presented in randomized order) among all species of living beings. Specifically, we asked: “What percentage of species consist of living beings that [feature]?” People’s answers were measured on a scale ranging from 0% to 100% in steps of 1%. The results are depicted in Figure 2.

We conducted Pearson’s χ^2 -tests to investigate which features yielded 100% as a response from significantly more than half the participants. For the features *organic material* ($\chi^2 = 14.286$, $p < 0.001$), *cells* ($\chi^2 = 9.143$, $p = 0.002$), and *evolution* ($\chi^2 = 5.143$, $p = 0.023$), a significant majority indicated that these features are universal, that is, that these features appear in living species with a frequency of 100%.

Whereas the purpose of Study 1 was to uncover the most salient features in our conception of life, Study 2 aimed to investigate whether some features are considered to be universal features of life. While organic composition, examined via the two features *cells* and *organic matter*, played hardly any role in establishing the salient features of the concept of life, it was considered to be a universal feature of life for around 2/3rds of the participants in Study 2. Consequently, organic composition is so far the best candidate to count as a necessary feature of the folk concept of life, given that only universal features can also be necessary features.

General Discussion

What does this all mean for the folk concept of life? Our experiments suggest that the concept of life is not as messy as has been suggested; rather, to most participants, a particular material constitution seems to form the essence of life. Additionally, functionings, such as nutrition and reproduction, but not material composition, turn out to be salient features commonly associated with living beings.

How can we explain our findings or at least organize them in a more illuminating way? We propose that the folk concept of life is a natural kind concept, elaborated as follows:

1. People think that life is a natural kind characterized by a certain essence.
2. They pick out the natural kind using salient features at the meso-level.
3. They think that the essence can be identified using scientific research. In addition to this, we propose:
4. People are inclined to take certain material building blocks (cells or a certain kind of matter) to be part of the essence of life.
5. They are open to revising this view on what is part of the essence of life (upon, for instance, the production of new findings by science that require such a revision)

This proposal is obviously inspired by a strand of thinking about natural kind terms prominently defended by Kripke (1980) and Putnam (1975), and others. The rough idea is that such a term picks out a natural kind not using its essence, but rather by giving examples, or typical observable features. The kind is supposed to have an essence that is left to investigation.

It is thus possible that some of the examples given do not belong to the kind. It may also turn out that some features used to pick out the kind are not instantiated by all members of the kind, or that they are not even typical of the kind.

Our proposal is closely related to recent work by Cleland (e.g. Cleland (2012, 2019)). She suspects that life is a natural kind, and she proposes to answer the question of what life is using a theoretical identity statement. Our proposal accords well with Cleland's main idea. If it is on the right track, then ordinary people assume there is an essence that living beings share, and they expect progress in finding the essence of life from science—as does Cleland. That said, there are differences between our proposal and Cleland's ideas. Most importantly, our proposal is about the folk concept of life, which Cleland isn't really interested in. While we propose that the use of the folk concept of life *presumes* it to be a natural kind, Cleland proposes that life is *in fact* a natural kind.

References

- [1] Cleland, C. (2012). Life without definition. *Synthese* 185, 125-144.
- [2] Cleland, C. E. (2019). *The quest for a universal theory of life: searching for life as we don't know it*. Cambridge: Cambridge University Press.
- [3] Kripke, S. A. (1980). *Naming and necessity*, Cambridge (MA), Harvard University Press.
- [4] Putnam, H. (1975). The meaning of 'meaning'. In K. Gunderson (Ed.), *Language, mind and knowledge: Minnesota studies in the philosophy of science* (Vol. VII, pp. 131-193). Minneapolis, MN.

Grounding minimal cognition in regulatory control

Leonardo Bich*
Universidad del Pais Vasco (UPV/EHU)

This work addresses the question of minimal cognition by investigating the relationship between cognitive capabilities the basic organisation of biological systems. More specifically, it provides a theoretical model of how a system can distinguish between specific features of its interaction with the environment and make decisions based on the evaluation of these distinctions: a fundamental requirement for the realization of minimal forms of cognition. Starting from the framework of biological autonomy it explores the deep connection between biological organisation and cognition by analysing the organisational requirements for the realisation of these minimally cognitive capabilities. It argues that they are grounded in the molecular domain and originate from mechanisms of biological regulation. All present-day living beings have this capacity of making decision based on distinctions made by regulatory mechanisms, let us just think of how bacteria respond adaptively to the composition (and variation in composition) of the environment by means of internal changes, such as the synthesis of different sets of enzymes necessary for metabolising different substances, or by modulating their movement in the environment according to gradients of concentrations, as in the case of chemotaxis.

The term cognition applies to diverse phenomena that figure in adaptive interactions between biological organisms and their environments. There is no agreement on whether and how it is possible to trace the boundaries between cognitive and non-cognitive (biological only) activities of biological systems. Are unicellular systems that are capable of chemotaxis, communication, vision already ‘cognitive’ (Lyon, 2015)? Or is a nervous system required, with the consequence that cognition is only found in animals (Christensen and Hooker, 2000; Barandiaran and Moreno, 2006)? Or is it restricted to organisms with a neocortex, or even to humans (Bayne et al., 2019)? How one answers these questions has implications ranging from how one characterizes the core mechanisms involved in cognition to how one selects model organisms for studying cognition.

Various theorists advance different criteria for attributing cognition. The dominant strategy has been to take human beings as the reference point and to focus on distinctively human cognitive activities. Lyon (2006) characterized this as the anthropogenic approach. She contrasts it with a biogenic approach, which locates cognition in the demands that

* leonardo.bich@ehu.es

organisms must meet in order to maintain themselves. Among those who embrace the biogenic approach, there are differences regarding how far back in phylogeny to identify cognition. Some, following Piaget (1967) and Maturana and Varela (1980), among others, identify cognition as an activity of all living organisms. This thesis, for which Heschl (1990) coined the name ‘Life = Cognition Thesis’, maintains that the dimension of a living organism interacting with its environment and modifying itself without losing its identity coincides with cognition. However, formulated in these terms, the thesis countenances all interactions between an organism and its environment, fails to identify the mechanisms underlying cognitive capabilities, and cannot distinguish cognition from mere causal interactions between organisms and their environments (Bich and Moreno, 2016).

This work embraces the view that cognitive activities are performed by all living organisms. Yet it identifies as minimally cognitive not all possible interactions between organism and environment, but only those activities resulting at least from the actions of regulatory control mechanisms over metabolic and agential activities. It focuses on one activity that is usually considered as fundamentally cognitive: decision-making. It grounds cognition, specifically decision-making, in an activity all organisms as autonomous systems must perform to keep themselves viable: controlling their basic (first-order) mechanisms. First-order mechanisms (called production or constitutive mechanisms) responsible for the basic activities carried out by living systems such as procuring resources from their environment, putting these resources to use to construct and repair the organism’s parts and moving through the environment. Regulatory mechanisms are higher-order mechanisms that control the activities of first-order mechanisms (Bich et al., 2016; Winning and Bechtel, 2018).

Given the variable nature of the environment and the continual degradation of the organism, first-order mechanisms must be regulated by mechanisms that select when a production is required and how it should be carried out in such a way as to maintain the viability of the system. This is done by regulatory mechanisms, which are endowed with sensory –effector capabilities, and operate on the basis of the evaluation of perturbations. To operate on first-order mechanisms, regulatory mechanisms need to make distinctions, i.e. to measure the state of the system and its environment, and evaluate possible actions. They are making decisions. By means of regulatory mechanisms, the system establishes some categories in the internal and external environment (sensory capability), and evaluates and employs them so as to select actions to functionally control its own internal dynamics (effector capability) in such a way that the system maintains its identity (Bich and Moreno, 2016; Bechtel and Bich, 2021).

In all organisms, these decisions are made by multiple different control mechanisms. In many cases, they employ internal models of features of the environment with which the organism must deal. Cognition, in the form of decision-making, is thus fundamental

to living systems which must control their first-order mechanisms. Decision-making will be discussed by analysing examples from bacteria that perform activities such as chemotaxis (Parkinson et al., 2015) and other organisms in which decisions are made on the basis of measurements made by regulatory mechanisms (Huang and Bechtel, 2020). But a more complex form of decision-making arises when organisms rely on internal models of the world. This may seem to go far beyond the capacities of bacteria, but some species of cyanobacteria rely on an internal model of the light–dark cycle (a circadian clock, see Mihalcescu et al., 2004) to regulate a host of activities (Bechtel 2011).

All present-day living beings have the capacity to make decisions based on regulatory mechanisms. They must control and coordinate first-order mechanisms so that they function in ways appropriate given the organism’s internal state and environmental conditions. Situating decision-making in the framework of control mechanisms allows to identify its core features and to underline the importance of studying it in simpler model organisms. A first advantage of this account is the fact that control mechanisms are ubiquitous in biological systems, thus radically expanding the range of model organisms in which we can study decision-making. The second advantage is that it allows to focus on relatively simple organisms such as bacteria, which are potent models to identify and analyse the core elements of a cognitive process such as decision-making and the regulatory mechanisms responsible for it, in line with the long tradition of research that emphasizes the continuity between biological and cognitive phenomena.

References

- [1] Barandiaran X, Moreno A. (2006). On what makes certain dynamical systems cognitive: a minimally cognitive organization program. *Adapt. Behav.* 14, 171-185.
- [2] Bayne T et al. (2019). What is cognition? *Curr. Biol.* 29, R608-R615.
- [3] Bechtel W. (2011). Representing time of day in circadian clocks. In *Knowledge and representation* (eds A Newen, A Bartels, E-M Jung), pp. 129-162. Palo Alto, CA: CSLI Publications.
- [4] Bechtel W, Bich L. (2021). Grounding cognition: heterarchical control mechanisms in biology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376 (1820), 20190751.
- [5] Bich L, Moreno A. (2016). The role of regulation in the origin and synthetic modeling of minimal cognition. *Biosystems* 148, 12-21.
- [6] Bich L, Mossio M, Ruiz-Mirazo K, Moreno A. (2016). Biological regulation: controlling the system from within. *Biol. Philos.* 31, 237-265.

- [7] Christensen WD, Hooker CA. (2000). An interactivist-constructivist approach to intelligence: self-directed anticipative learning. *Philos. Psychol.* 13, 5-45.
- [8] Heschl A. (1990). L=C a simple equation with astonishing consequences. *J. Theor. Biol.* 145, 13-40.
- [9] Huang LT, Bechtel W. (2020). A phylogenetic perspective on distributed decision-making. In Proc.42nd Annu. *Meeting Cogn. Sci. Soc.* (eds S Denison, M Mack, Y Xu, BC Armstrong), pp. 3191-3197. Seattle, WA: Cognitive Science Society.
- [10] Lyon P. (2006). The biogenic approach to cognition. *Cogn. Process.* 7, 11-29. (doi:10.1007/s10339-005-0016-8)
- [11] Lyon P. (2015). The cognitive cell: bacterial behaviour reconsidered. *Front. Microbiol.* 6, 264.
- [12] Maturana HR, Varela FJ. (1980). *Autopoiesis and cognition: the realization of the living*. Dordrecht: Reidel.
- [13] Mihalcescu I, Hsing W, Leibler S. (2004). Resilient circadian oscillator revealed in individual cyanobacteria. *Nature* 430, 81-85.
- [14] Moreno A, Mossio M. (2015). *Biological autonomy: a philosophical and theoretical inquiry*. Dordrecht: Springer.
- [15] Parkinson JS, Hazelbauer GL, Falke JJ. (2015). Signaling and sensory adaptation in *Escherichia coli* chemoreceptors: 2015 update. *Trends Microbiol.* 23, 257-266.
- [16] Piaget J. (1967). *Biologie et connaissance*. Paris: Gallimard.
- [17] Winning J, Bechtel W. (2018). Rethinking causality in neural mechanisms: constraints and control. *Minds Mach.* 28, 287-310.

On particle disappearance in classical mechanics

Amaia Corral-Villate*
University of the Basque Country

The context of this talk is provided by the particle disappearance conclusions claimed in some recent discussions on classical mechanical particle configurations, namely, those by Alper & Bridger (2002), Pérez Laraudogoitia (1998) and Shackel (2018).

The objective of the talk is to clarify the basis on which particle disappearance may justifiably be claimed, and consequently determine whether those three particular configurations fulfil the required characteristics. The key to the diagnosis of these requirements and the consequent disappearance result is shown by Corral-Villate (2020) to be given by the application of a fundamental continuity condition compatible with the classical mechanical principle of mass conservation.

The results of this analysis lead to conclude that because both the discussions by Alper & Bridger (2002) and Shackel (2018) imply the violation of this fundamental principle, in neither of these configurations is the particle disappearance claim justified. In contrast to this, in the case of the system by Pérez Laraudogoitia (1998) the required principle is fulfilled and thus the particle extinction conclusion is justified.

References

- [1] Corral-Villate, A. (2020). On Shackel's nothing from infinity paradox. *European Journal for Philosophy of Science*, 10 (3), 1-13.
- [2] Pérez Laraudogoitia, J., Bridger, M. & Alper, J. (2002). Two Ways of Looking at a Newtonian Supertask. *Synthese*, 131 (2), 173-189.
- [3] Pérez Laraudogoitia, J. (1998). Infinity machines and creation ex nihilo. *Synthese*, 115 (2), 259-265.
- [4] Shackel, N. (2018). The infinity from nothing paradox and the immovable object meets the irresistible force. *European Journal for Philosophy of Science*, 8 (3), 417-433.

* amaiacorral@yahoo.co.uk

Individualidad biológica y viviparidad: una investigación filosófica desde las transiciones evolutivas y la evo-devo

David Cortés García*
Universidad del País Vasco, UPV/EHU

Arantza Etxeberria Agiriano**
Universidad del País Vasco, UPV/EHU

Resumen extendido

En este trabajo examinamos algunos de los aspectos filosóficos más relevantes involucrados en el cambio de modo de reproducción de la oviparidad a la viviparidad, en relación con la aparición de nuevas ontologías. Para ello, examinamos la aparición de esta novedad evolutiva desde dos puntos de vista: por un lado, desde el estudio del origen de las novedades en la evo-devo y, por otro lado, desde la teoría de las transiciones evolutivas. Nuestra hipótesis es que en la evolución de las distintas formas de reproducción podemos observar la emergencia de nuevas clases de individualidad, aunque pueda resultar controvertido que estas constituyan transiciones evolutivas en el marco estándar de la teoría de las transiciones evolutivas.

Evolución de la viviparidad

La reproducción vivípara o viviparidad es entendida como un patrón reproductivo caracterizado por la retención de los embriones en el interior del cuerpo de la madre, de manera que la descendencia es liberada al entorno en forma de larva o juvenil, en contraste con la oviparidad, caracterizada por la puesta de huevos. La viviparidad está asociada con el suministro directo y constante de nutrientes al embrión, denominado matrotrofia. Aquí, distinguimos tres características fundamentales de la evolución de la viviparidad y la matrotrofia: (i) su *prevalencia* a lo largo del reino animal, estando presente en 20 de los 32 filos animales (Ostrovsky et al., 2016); (ii) su *carácter homoplásico*, evidenciado por la existencia de más de 150 orígenes independientes tan solo en vertebrados (Blackburn, 2015); y (iii) su *realizabilidad múltiple*, determinada por el amplio espectro de estrategias diferentes a través de las cuales puede lograrse la viviparidad, desde la absorción de nutrientes (por endocitosis, o mediada por estructuras como la placenta) hasta la ingestión

* davidcortescgarcia.dcg@gmail.com

** arantza.etxeberrria@ehu.eus

de tejido materno, huevos no fertilizados u otros embriones (Blackburn, 2015; Dopazo & Alberch, 1994; Fusco & Minelli, 2019).

Novedades desde la evo-devo

Las novedades involucradas en la evolución de la viviparidad –concretamente la internalización del desarrollo embrionario y la separación del entorno– se han estudiado desde los enfoques de la macroevolución y de la evo-devo. La primera aproximación entiende que la evolución de la viviparidad refleja una tendencia evolutiva más amplia hacia un incremento en la autonomía, de manera que las influencias de las fluctuaciones del entorno sobre el organismo individual se ven gradualmente reducidas (Rosslenbroich, 2014). La evolución de la viviparidad en mamíferos euterios y, consecuentemente, del embarazo, constituye un ejemplo particularmente destacado de esta tendencia. Por otro lado, desde la evo-devo se ha propuesto que la implantación del embrión en la pared uterina en euterios –que involucra la aparición de nuevos tipos celulares e innovaciones funcionales– evolucionó a partir de la respuesta celular al estrés (Erkenbrack et al., 2018). Esta respuesta inmunológica, inducida por alteraciones en el entorno tales como la escasez de alimentos o la exposición a tóxicos, es *internalizada*, dando lugar a un proceso fisiológicamente regulado (Wagner et al., 2019). Así, las rutas de la respuesta al estrés ambiental pasan a estar reguladas de forma interna, de manera que la consecuente respuesta inflamatoria induce la implantación del embrión, permitiendo la evolución de la placentación invasiva. A raíz de estos hallazgos, se ha argumentado que el embarazo es una novedad relacional altamente cooperativa y de carácter sistémico, cuya evolución implica la emergencia de un *individuo histórico*: la hembra embarazada (Nuño de la Rosa et al., 2021).

En relación con esto, nos planteamos la cuestión acerca de si la evolución de la viviparidad podría entenderse como una forma de transición evolutiva en la que madre y embrión(es) constituyen, de forma temporal, un único individuo. Su formación y desvanecimiento están marcados por dos eventos inflamatorios (la implantación y el parto, respectivamente), que acotan temporalmente el individuo emergente de la hembra embarazada (Nuño de la Rosa et al., 2021).

Transiciones evolutivas

En el enfoque clásico de las transiciones evolutivas, iniciado por John Maynard Smith y Eörs Szathmáry, las transiciones evolutivas se producen cuando un conjunto de unidades biológicas más pequeñas, que originalmente eran capaces de sobrevivir y reproducirse por sí solas, se agrupan en una sola unidad mayor, generando un nuevo nivel de organización en la jerarquía biológica (Okasha, 2019). Estas transiciones dan cuenta de los orígenes de nuevas unidades en la evolución tales como los cromosomas, los eucariotas,

los seres multicelulares o los grupos sociales. Las transiciones evolutivas se caracterizan por que “las entidades que eran capaces de *replicarse* de forma independiente antes de la transición sólo pueden *replicarse* como parte de un todo mayor después de ella” (Maynard Smith & Szathmáry, 1995, p. 6, cursiva nuestra). Una vez formada la entidad a un nivel de organización superior, los componentes pierden la capacidad de auto-replicación, de forma que el *fitness* de los componentes está supeditado al *fitness* de la entidad del nivel superior; de lo contrario, se produciría la disrupción del sistema. Posteriormente, y con el objetivo de integrar en este marco dos aspectos fundamentales como son la herencia y el desarrollo, James Griesemer propone que las transiciones evolutivas implican que “las entidades que eran capaces de *reproducirse* de forma independiente antes de la transición sólo pueden *reproducirse* como parte de un todo mayor después de ella” (Griesemer, 2000, p. 79, cursiva nuestra). Por tanto, las transiciones evolutivas implican cambios importantes en los modos de reproducción y están asociadas con la aparición de nuevos tipos de individuos. Por otro lado, David Queller (2000) distingue las transiciones *fraternales* de las *igualitarias* en evolución: las primeras se refieren a la aparición de nuevas unidades mediante la asociación de entidades homogéneas, como la agrupación de organismos unicelulares que generan individuos multicelulares, o la formación de colonias por asociación de individuos; mientras que las segundas se refieren a la fusión de entidades heterogéneas, como el acoplamiento de genes que genera el cromosoma, o la asociación de células procariotas que da lugar a la aparición de la célula eucariota.

Volviendo al caso de la viviparidad en euterios, parece controvertido que la evolución del embarazo se trate, de forma estricta, de una transición evolutiva en el sentido estándar de esta teoría. Sin embargo, no es menor el hecho de que en el embarazo se forme una nueva forma de individualidad, por lo que cabe explorar de forma concienzuda los aspectos ontológicos involucrados en esta innovación evolutiva.

Discusión

La teoría de las transiciones evolutivas y el enfoque sobre las novedades ofrecido por la evo-devo comprometen ontologías aparentemente irreconciliables sobre las unidades evolutivas y cómo estas interactúan y se transforman. Mientras que la primera aproximación ofrece un marco sobre la transición de individualidades basada en la teoría de la selección multinivel y con el foco en el *fitness* relativo de cada una de las unidades o entidades involucradas, para la perspectiva de la evo-devo la distinción e individuación de unidades evolutivas (tales como homologías o planes corporales) se basa en la evolución y estabilización histórica de elementos (órganos y/o individuos) homólogos, caracterizados por sus mecanismos de desarrollo, su persistencia a lo largo del tiempo evolutivo y por tener una historia evolutiva propia.

La cuestión sobre si ambos enfoques son, efectivamente, opuestos e irreconciliables o si, más bien, podrían entenderse como compatibles e incluso complementarios en el marco de un pluralismo ontológico, es fundamental para entender la ontología evolutiva de la reproducción y poder dar respuesta a cuestiones como la planteada aquí acerca de la conveniencia de considerar la evolución de la viviparidad en euterios como una transición evolutiva.

Referencias

- [1] Blackburn, D. G. (2015). Evolution of vertebrate viviparity and specializations for fetal nutrition: A quantitative and qualitative analysis: Viviparity and Fetal Nutrition. *Journal of Morphology*, 276(8), 961-990. <https://doi.org/10.1002/jmor.20272>
- [2] Dopazo, H., & Alberch, P. (1994). Preliminary results on optional viviparity and intrauterine siblicide in *Salamandra salamandra* populations from Northern Spain. *Mertensiella*, 4, 125-137.
- [3] Erkenbrack, E. M., Maziarz, J. D., Griffith, O. W., Liang, C., Chavan, A. R., Nnamani, M. C., & Wagner, G. P. (2018). The mammalian decidual cell evolved from a cellular stress response. *PLOS Biology*, 16(8), e2005594. <https://doi.org/10.1371/journal.pbio.2005594>
- [4] Fusco, G., & Minelli, A. (2019). *The Biology of Reproduction*. Cambridge University Press. <https://doi.org/10.1017/9781108758970>
- [5] Griesemer, J. (2000). The Units of Evolutionary Transition. *Selection*, 1(1-3), 67-80. <https://doi.org/10.1556/Select.1.2000.1-3.7>
- [6] Maynard Smith, J., & Szathmáry, E. (1995). *The major transitions in evolution*. Oxford Univ. Press.
- [7] Nuño de la Rosa, L., Pavličev, M., & Etxeberria, A. (2021). Pregnant Females as Historical Individuals: An Insight From the Philosophy of Evo-Devo. *Frontiers in Psychology*, 11. <https://doi.org/10.3389/fpsyg.2020.572106>
- [8] Okasha, S. (2019). *Philosophy of Biology*. Oxford University Press.
- [9] Ostrovsky, A. N., Lidgard, S., Gordon, D. P., Schwaha, T., Genikhovich, G., & Ereskovsky, A. V. (2016). Matrotrophy and placentation in invertebrates: A new paradigm. *Biological Reviews*, 91(3), 673-711. <https://doi.org/10.1111/brv.12189>
- [10] Queller, D. C. (2000). Relatedness and the fraternal major transitions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 355(1403), 1647-1655.

- [11] Rosslénbroich, B. (2014). *On the Origin of Autonomy: A New Look at the Major Transitions in Evolution*.
- [12] Springer International Publishing. <https://doi.org/10.1007/978-3-319-04141-4>
- [13] Wagner, G. P., Erkenbrack, E. M., & Love, A. C. (2019). Stress-Induced Evolutionary Innovation: A Mechanism for the Origin of Cell Types. *BioEssays*, 1800188. <https://doi.org/10.1002/bies.201800188>

El experimento científico por simulación y el problema de la materialidad

Elena Denia*
Universidad de Valencia

Desde el punto de vista filosófico, las simulaciones informáticas comparten aspectos de la teorización y de la experimentación en ciencia. En el primer caso, puede decirse que se trabaja con los modelos que articulan el programa en la computadora; en el segundo, se trabaja con datos empíricos. Pero la relación entre ambas actividades es íntima y su propia naturaleza impide estudiarlas por separado. Cuando reflexionamos acerca de la experimentación tradicional –la caracterizada por Hacking, Franklin y otros autores– una pregunta fundamental que surge en este contexto es: ¿las simulaciones son experimentos? Y si no lo son, ¿qué características esenciales, si las hay, distinguen las simulaciones y los experimentos? O, en cualquiera de los dos casos, ¿permiten revelar contenido empírico de las teorías?

Recurrimos tanto a la simulación como a la experimentación cuando no tenemos una descripción completa del comportamiento de un sistema objetivo que queremos estudiar. La simulación a veces presenta un carácter mimético y produce imágenes similares a las que vemos en el mundo (Winsberg, 2010), y otras proporciona datos de salida dispuestos en un sistema de representación con el que estamos familiarizados; en ambos casos puede revelarnos información cualitativa sobre el mundo. En este sentido, puede jugar un papel relevante en la producción de conocimiento científico. En concreto, algunos autores han propuesto que la simulación se encuentra entre los dominios experimental y teórico, sin pertenecer enteramente a ninguno de los dos, pero estando a medio camino entre ambos (Galison, 1996). Otros académicos, en cambio, sugieren que la simulación trae consigo una nueva comprensión filosófica de la ontología –por ejemplo, Giere (2009) considera que “simulación por ordenador es un fenómeno cualitativamente nuevo en la práctica de la ciencia”–.

El debate puede abordar atención a la distinción entre simulación y experimento tradicional. En los experimentos tradicionales, los científicos tratan de aprender sobre un sistema y esperan poder hacer generalizaciones aplicables a otros sistemas, y para justificarlas suelen apelar a similitudes materiales (Parker, 2009). Dado que los estudios por simulación también comparten este propósito, cabe preguntarse si tales generalizaciones pueden justificarse indistintamente tanto a través del plano material como del abstracto.

* elenismou@gmail.com

Por ello, una consideración central en el debate que se ha discutido intensamente es si la “materialidad” es una característica intrínseca de los experimentos.

Nótese que la materialidad juega un papel esencial en los argumentos filosóficos cuando se pretende evaluar la capacidad epistémica por parte de los filósofos. Al respecto, algunos consideran que la simulación es no material –llegando a afirmar que esa carencia empobrece su capacidad epistémica–, mientras que otros señalan que el sistema está programado en una computadora digital, que es física. Una consideración central en el debate es la propuesta por Giere (2009), quien apunta que en el experimento tradicional existe una interacción causal con el sistema objetivo, y por ello, según se ajuste o no a este último, el resultado epistemológico tendrá mayor o menor confianza. Por su parte, Morgan señala que el experimento material ocupa una posición epistémica privilegiada porque la validez externa en la simulación es más complicada. Sin embargo, también se ha aducido que los dispositivos computacionales “extienden nuestras limitadas potencias de cálculo de manera similar a las formas en que los instrumentos empíricos extienden nuestras limitadas potencias de observación” (Humphreys, 2004: 116). En particular, a través de la simulación informática es posible ampliar las matemáticas manejables y aumentar el rango de los objetos epistemológicamente accesibles (Frigg y Reiss, 2009). En este sentido, la potencia de cálculo podría proporcionar superioridad epistémica.

Una forma de abordar el problema es reparar en que la importancia del experimento radica en las “similaridades relevantes” con el “sistema objetivo”, ya sean materiales y/o formales. En concreto, en el experimento tradicional, el objeto y el objetivo presentan una similitud material, mientras que en la simulación esa similitud es abstracta o formal, y así “en los experimentos genuinos, las mismas causas ‘materiales’ están en juego en los sistemas experimental y objetivo, mientras que en las simulaciones existe una correspondencia meramente formal entre los sistemas de simulación y objetivo” (Parker, 2009: 484).

Tras un análisis de las distintas posturas, defiende que no hay una superioridad epistémica que, en términos generales, se pueda atribuir automáticamente a la materialidad, de la misma forma que tampoco puede atribuirse esa superioridad a la potencia de cálculo; a no ser que examinemos casos particulares con propósitos específicos. Mi postura es que tanto la materialidad como la capacidad de cálculo proporcionan potencia epistémica, pero no una superioridad epistémica de la una sobre la otra, puesto que actúan en planos incomparables –el abstracto y el material– y cada diseño experimental exhibe sus propias características relevantes.

En particular, cuando solo trabajamos con sistemas formales, para determinar si los criterios son internos a la representación misma, siguiendo a Hacking (1983), “entra en juego la metafísica para tratar de distinguir los buenos sistemas de representación de los malos”, una observación que muestra la complejidad al debate. Si como afirma Parker, “la relevancia es una función de las preguntas particulares que se hacen sobre el sistema

objetivo” (Parker, 2009: 495), los dominios o espacios de solución para responder ciertas preguntas pueden operar únicamente en el plano abstracto. Queda señalar que la dificultad de desentrañar cuáles son las similitudes relevantes entre el sistema objetivo y el objeto, no obstante, es una cuestión aparte que requiere un análisis más profundo.

Bibliografía

- [1] Galison, P. (1996). Computer simulations and the trading zone. In P. Galison & D. J. Stump (Eds.), *The Disunity of Science: Boundaries, Contexts, and Power* (pp. 118-157): Stanford University Press.
- [2] Giere, R. N. (2009). Is computer simulation changing the face of experimentation? *Philosophical Studies*, 143(1), 59-62. doi:10.1007/s11098-008-9314-1
- [3] Hacking, I. (1983). *Representar e intervenir*. México DF: Paidós.
- [4] Humphreys, P. (2004). *Extending ourselves: Computational science, empiricism, and scientific method*: Oxford University Press.
- [5] Parker, W. S. (2009). Does Matter Really Matter? Computer Simulations, Experiments, and Materiality. *Synthese*, 169(3), 483-496. Retrieved from www.jstor.org/stable/40271305
- [6] Winsberg, E. (2010). *Science in the age of computer simulation*: University of Chicago Press.

Attribution methods in climate science: exploring the relevant differences

Laura García Portela*
University of Graz

We know that climate change is a human-induced phenomenon and the product of greenhouse gas emissions-generating activities. We are less certain about the link between particular extreme weather events (EWEs) and anthropogenic climate change forcing since EWEs would occur even in a pristine climate. However, this link seems to be of particular interest, given the already-occurring and projected impacts of EWEs on individuals and property. Attribution of EWEs to anthropogenic forcing has become a topic of major scientific and social concern. At stake is the use of scientific evidence for pursuing certain socially relevant policies (such as adaptation planning) or, even, for advancing certain justice claims (such as those related to compensation for loss and damage).

Scientists have contributed to those pressing social demands by proposing different attribution methodologies: the risk-based approach and the storyline approach. There is a live debate between the convenience of relying on one or the other approach for attribution studies. Recently, Elisabeth Lloyd, Naomi Oreskes and Eric Winsberg (Lloyd and Oreskes 2018; 2019; Winsberg, Oreskes, and Lloyd 2020) have argued that the controversy should be seen through the lenses of error-type propensity and that the choice between those alternatives should therefore depend on the type of errors policy and decision-makers seek to avoid. In the following, I explain the two methodologies and argue that the analysis provided by those philosophers is wrongheaded.

The risk-based approach has been so far the conventional methodology in attributing EWEs to human forcing (Allen 2012). This methodology takes a certain event as a token of a class of EWEs and asks the following research question: How much did anthropogenic climate change increase the probability or risk of a specific type of event? Answering this question requires comparing the probability (P_1) or risk of a specific class of event in a world affected by anthropogenic climate change (actual world) and the probability (P_0) of such a type of event in a world without climate change (hypothetical counterfactual world). Once we have both probabilities, P_1 and P_0 , the result is expressed as a probability ratio. Usually, this operation is represented with a Fraction of Attributable Risk (FAR), defined as $FAR = 1 - (p_0/p_1)$. The FAR is interpreted as the fraction of the likelihood of an event that is attributable to the external forcing. Hence, FAR leads to probabilistic causal

* laura.garciaportela@gmail.com

claims, such as: an event of that class was 70% more likely to occur in a world with climate change than in a world without climate change.

The storyline approach emerges from scepticism concerning applying probability assessments to a certain type of EWEs. To explain this scepticism, proponents of the storyline approach distinguish between the dynamical and thermodynamical climate variables involved in a EWE (Trenberth, Fasullo, and Shepherd 2015; Shepherd et al. 2018). Advocators of the storyline approach argue that identifying human contribution to changes in dynamical climate variables is very challenging and it often delivers unreliable or inconclusive results. For that reason, they take the dynamically-driven climate variables of an event as a given constraint and then ask about the contribution of human forcing to the event's thermodynamical climate variables. In that way, they obtained answers to the attribution question under the condition that the dynamics would remain the same. They shift the research question into one about the event's magnitude or severity. That is, the relevant research question is: 'how much has climate change increased the magnitude of this event'?

According to Lloyd, Oreskes and Winsberg, the risk-based approach favours type II errors (false negatives), whereas the storyline approach favours type I errors (false positives). However, I argue here that this categorization of attribution methodologies should be rejected for the following reasons. First, it is doubtful that, overall, the risk-based approach can be considered prone to type II errors since the problem is that climate models do not realistically simulate changes in the dynamics (Woollings 2010; Shepherd 2014). The lack of robustness due to model errors make it difficult to trust the results concerning events mostly driven by dynamically-driven climate variables but that does not mean that the methodology is prone to false negatives.

Second, there are two reasons to cast doubt into the identification of the storyline approach with a propensity to type I errors. As Winsberg, Lloyd and Oreskes claim themselves, the storyline approach is simply asking a conditional question, such as: taking the dynamically-driven climate variables as fixed, how has climate change affected the magnitude of the event? Claiming that this approach is prone to false positives is misunderstanding the conditional nature of the methodology. Moreover, by assuming that the Storyline Approach is prone to false-positive results, these scholars seem to believe that the dynamics of the atmosphere that are taken as a given yield results affirming that climate change had more influence on the overall event than it actually had. However, those actual changes in the dynamics could have also made the event more likely or severe. If that were the case, the approach would be underestimating the role of human forcing, thereby falling into false-negative results. Given that we do not have a reliable representation of the actual changes in the dynamics due to climate change, we cannot know whether

presupposing certain parameters for the dynamics of the event, as the Storyline Approach does, delivers false positive or false negative results.

Attribution studies are enormously important for adaptation planning and compensatory demands for the effects of climate change. For that reason, it is important to investigate the relevant differences between different methodologies and their consequences for decision-making in the context of climate change. As I have argued here, we should be wary of differentiating these two existing methodologies in terms of their propensity to certain error-types and, thus, such differentiation should play no role in decision-making concerning climate change measures.

References

- [1] Allen, Myles. (2012). ‘The Scientific Basis for Climate Change Liability’. In *Climate Change Liability: Transnational Law and Practice*, edited by Richard Lord, 8-22. Cambridge; New York: Cambridge University Press.
- [2] Lloyd, Elisabeth, and Naomi Oreskes. (2018). ‘Climate Change Attribution: When Is It Appropriate to Accept New Methods?’ *Earth’s Future* 6 (3): 311-325.
- [3] — (2019). ‘Climate Change Attribution: When Does It Makes Sense to Add Methods?’ *Epistemology and Philosophy of Science* 56 (1): 185-201.
- [4] Shepherd, Theodore G. (2014). ‘Atmospheric Circulation as a Source of Uncertainty in Climate Change Projections’. *Nature Geoscience* 7 (10): 703-708.
- [5] Shepherd, Theodore G., Emily Boyd, Raphael A. Calel, Sandra C. Chapman, Suraje Dessai, Ioana M. Dima-West, Hayley J. Fowler, et al. (2018). ‘Storylines: An Alternative Approach to Representing Uncertainty in Physical Aspects of Climate Change’. *Climatic Change* 151 (3-4): 555-571.
- [6] Trenberth, Kevin E., John T. Fasullo, and Theodore G. Shepherd. (2015). ‘Attribution of Climate Extreme Events’. *Nature Climate Change* 5 (8): 725-730.
- [7] Winsberg, Eric, Naomi Oreskes, and Elisabeth Lloyd. (2020). ‘Severe Weather Event Attribution: Why Values Won’t Go Away’. *Studies in History and Philosophy of Science Part A* 84 (December): 142-149.
- [8] Woollings, Tim. (2010). ‘Dynamical Influences on European Climate: An Uncertain Future’. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 368 (1924): 3733-3756.

Estimating closeness to the truth for scientific laws

Alfonso García-Lapeña*
University of Barcelona, Logos, BIAP

Truthlikeness is a property of a theory or a proposition that represents its ‘closeness’, ‘similarity’ or ‘likeness’ to the truth. Since Popper’s (1963) failure to provide a satisfactory definition of the concept, the notion has been a topic of intense discussion by philosophers of science and logicians, particularly as it seems an indispensable element for a plausible formulation of scientific realism and scientific progress.

As with many other philosophical concepts, we must clearly distinguish between: (a) the logical or semantic problem (how it is meaningful to claim that T_2 is more truthlike than T_1) and (b) the epistemological problem (how, given some evidence, it is rational to claim that T_2 is more truthlike than T_1).

In the present paper we will deal with the epistemological problem of truthlikeness for quantitative deterministic laws (QDL henceforth), providing an *estimation measure* of the degree of truthlikeness for QDL based on Niniiluoto’s (1987) semantic proposal and García-Lapeña’s (2021a, b) expanded version. Finally, we will apply our proposal to a real case study, estimating the degrees of truthlikeness of nine historical thermodynamic gas laws for nitrogen.

According to the similarity approach, ‘theory A is more truthlike than theory B ’ means that: given a space S of possible answers for a cognitive problem where the theories A , B and the truth in question T are represented; and given some appropriate similarity metric d for S ; then $d(A, T) < d(B, T)$. Applying this basic schema to the case of QDL, the possible behaviours of a phenomenon or a system can be represented in the *state space*, which is a mathematical abstract space composed by the quantities that characterize the system (for example, in classical thermodynamics the state of a gas is fixed by pressure, volume and temperature; therefore, the space of possible answers or states can be taken to be \mathbb{R}^3). A QDL can be represented as a function in the state space. Then, given a state space and a set of possible functions, what we need to define truthlikeness for QDL is an appropriate metric d which defines the distance (in terms of similarity) between functions.

* aglapena@gmail.com

Niiniluoto (1987, 2018) proposes to define the distance between a deterministic law X and the true deterministic law T (the true “connexion” between the quantities) by the *Minkowski* or L_p metric for functions ($p \geq 1$):

$$d(X, T) = \left(\int |f^X(x) - f^T(x)|^p dx \right)^{\frac{1}{p}}$$

Particularly, Niiniluoto advocates for $p = 1$ and for $p = 2$, which produce the well-known Manhattan distance ($d^{ma}(X, T)$) and Euclidean distance ($d^{eu}(X, T)$) between functions. Both represent one of the most natural and intuitive ways of measuring the distances between points and functions, being the L_p metric one of the basic constituents of functional analysis (Deza, 2013).

Niiniluoto’s proposal, however, presents unintuitive results in a number of cases. Consider the following case (García-Lapeña, 2021a) depicted in Figure 1:

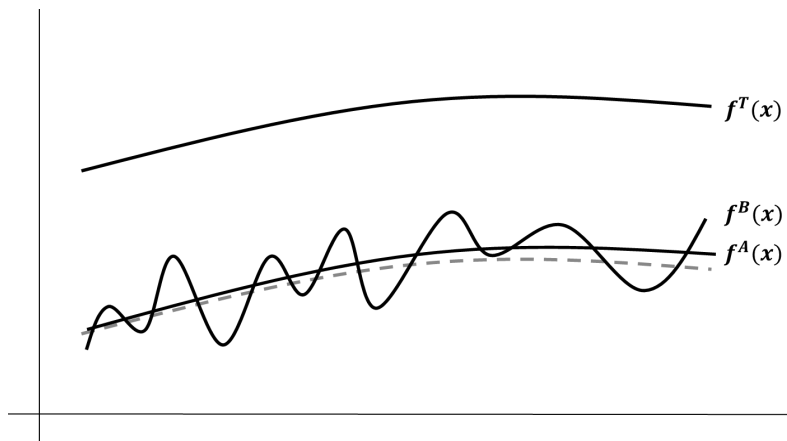


Figure 1

Consider that $f^A(x)$ and $f^B(x)$ are such that $d(A, T) = d(B, T)$, that is to say, that A and B are at the same distance from the truth (where distance is defined according to $d^{ma}(X, T)$ or $d^{eu}(X, T)$). In that situation, Niiniluoto’s proposal yields the result that both laws are equally truthlike. However, García-Lapeña take it as strongly intuitive that $f^A(x)$ appears to be more similar to $f^T(x)$ than $f^B(x)$. Law $f^A(x)$ clearly seems to be closer to how the world is or more similar to the actual behaviour of the system than $f^B(x)$. Therefore, a satisfactory definition of truthlikeness should give the result that $A >_t B$.

García-Lapeña argues that the underlying problem with Niiniluoto’s proposal lies in the fact that $d^{ma}(X, T)$ or $d^{eu}(X, T)$ measure the accuracy of a law, as they compare each predicted value to the corresponding true value. However, Figure 1 shows that two laws may be equally accurate and still one may imply more true or truthlike consequences,

behaviours or facts about the system than the other. Therefore, that the “shape” of a law seems to represent an additional factor which should play a role in defining truthlikeness. He calls this additional factor ‘nomicity’ and proposes to define the distance $d^{an}(X, T)$ between a law X and the true law T as:

$$d^{an}(X, T) = F(\text{accuracy}, \text{nomicity})$$

García-Lapeña further proposes that nomicity can be captured by the distance between the derivative functions. Opting in both cases for the Euclidean distance:

$$d^{an}(X, T) = F(d^{eu}(X, T), d^{eu}(X', T'))$$

The next step is to combine both factors into a single function in order to make comparative judgements and to obtain the numerical degree of truthlikeness of a law. García-Lapeña (2021b) argues for the following combination:

$$d^{an}(X, T) = (d^{eu}(X, T) + 1)(d^{eu}(X', T') + 1) - 1$$

Which can be then normalized and used to define the degree of truthlikeness of a QDL as:

$$Tr(X) = \frac{1}{(1 + d^{eu}(X, T))} \frac{1}{(1 + d^{eu}(X', T'))}$$

Based on this proposal, we elaborate a possible solution to the *epistemological* problem and test it in a case study regarding the evolution of nine thermodynamic gas laws.

The main difficulty in applying $d^{an}(X, T)$ to real cases lies on the fact that we don’t know the *truth*, we don’t know *how* the true law T is. What we generally have is a series of empirical observations which correspond to a set of points in a *state-space*.

Suppose a simple case where the target system is represented according to two quantities X and Y and that we have a set of n empirical observations $[(x_1^T, y_1^T), \dots, (x_n^T, y_n^T)]$. Although technically these values come with a margin of error ε , for our purposes we can make the simplification of assuming (x_i^T, y_j^T) as the “true values”.

The *estimated degree of accuracy* of a law X , $E[d^{eu}(X, T)]$, can be calculated by comparing, for each empirical point, the true value y_i^T with the predicted value y_i^X :

$$E[d^{eu}(X, T)] = \left(\frac{x_n^T - x_1^T}{n} * \sum_1^n (y_i^T - y_i^X)^2 \right)^{\frac{1}{2}}$$

Factor $\frac{x_n^T - x_1^T}{n}$ is introduced to approximate the summation to the *Riemann sum* of interval $x_n^T - x_1^T$, which in turn is an approximation to the definite integral of the interval (whose value would be the theoretical degree of accuracy of law X).

Nomicity, however, is harder to *estimate*. In order to calculate the derivative on a point the values of the point are not enough, we need to know the function itself. As we don't know how the "true function" is, we can't directly calculate the values of the derivatives of the empirical points.

The only way to proceed, then, is to rationally estimate by some procedure the value of the derivative of the true function for each empirical point. Given a set of known data points there are many statistical methodologies for interpolating the function that best fits the set of points. In a sense there is no "best" interpolation method, but each presents advantages and disadvantages. What we need is a rational measure of the "slope" of the real function when it "passes" through the known empirical points.

For this purpose, we think that *cubic splines* stand in a privileged position. A *cubic spline interpolation* results in a piecewise function with each of its pieces being a three order polynomial. What makes cubic splines philosophically interesting for our purposes is that they instantiate the *minimum curvature property* (Holladay, 1957), which can be roughly stated as follows: among all possible curves that can pass through a set of given points, cubic splines represent the path which presents minimum curvature (Leslie *et al.*, 1996). Roughly, they can be understood as defining the "shortest" plus "smoothest" path between a set of known points. Philosophically, we take *cubic splines* as potential good instantiations of the metaphysical principle of *simplicity* in nature, such that *cubic splines* estimate the "most probable path" of the *True law* in the proximities of each empirical point, and so they enable us to estimate the "most probable value" of the *True derivative* for each empirical point.

Therefore, the *estimated degree of nomicity* of a law X , $E[d^{eu}(X', T')]$, can be calculated by comparing, for each empirical point, the value of the derivative of the spline ($y_i^{S'}$) with the value of the derivative of law X ($y_i^{X'}$):

$$E[d^{eu}(X', T')] = \left(\frac{x_n^T - x_1^T}{n} * \sum_1^n (y_i^{S'} - y_i^{X'})^2 \right)^{\frac{1}{2}}$$

We then conduct a historical analysis of the development of classical thermodynamic laws, which concludes in two main families of gas laws (Sengers 2000) and nine relevant

laws: *Ideal Gas*, *Van der Waals*, *Dieterici*, *Berthelot*, *Redlich-Kwong*, *Peng-Robinson*, *Beattie-Bridgeman* and *Benedict-Webb-Rubin*.

We proceed to calculate the *estimated accuracy*, *estimated nomicity* and *estimated truthlikeness* of the nine historical gas laws regarding a sample of 421 empirical observations of nitrogen in gas state. The results match with the standard scientific considerations about the laws and with the evolution of the historical episodes, giving plausibility to our proposal.

References

- [1] García-Lapeña, A. (2021a). Truthlikeness for quantitative deterministic laws. *The British Journal for the Philosophy of Science*. Forthcoming.
- [2] García-Lapeña, A. (2021b). Truthlikeness for probabilistic laws. *Synthese*. Forthcoming.
- [3] Holladay, J. (1957). A Smoothest Curve Approximation. *Mathematical Tables and Other Aids to Computation*, 11(60).
- [4] Leslie F. *et al.* (1996) *Advances in Morphometrics*. Springer Science & Business Media.
- [5] Niiniluoto, I. (1987). *Truthlikeness*. Dordrecht: Reidel.
- [6] Niiniluoto, I. (2018). *Truth-Seeking by Abduction*. Springer.
- [7] Popper, K. [1963]: *Conjectures and refutations*, London: Routledge and Kegan Paul.
- [8] Sengers, J. *et al.* (2000). Equations of state for fluids and fluid mixtures. Part 1. *Experimental Thermodynamics* v. 5. Elsevier.

La narración del paciente como “evidencia epistemológica” en medicina

Natividad Garrido Rodríguez*
Universidad de La Laguna

Dentro del espacio sanitario proliferan una multiplicidad de narrativas que manifiestan los distintos puntos de vista de la experiencia de la enfermedad. Sin embargo, no todas estas narrativas tienen un papel dentro del quehacer médico, ya que el tipo de relato que se prioriza y al que se le da valor de verdad y de objetividad es la historia clínica escrita por el médico. Una historia clínica que se construye tomando como punto de partida la narración del paciente pero, bajo la aureola de que ésta ha de ser lo más objetivamente posible para poder realizar un diagnóstico. La interacción en este momento suele ser fría y un tanto angustiada, con preguntas cerradas (sí/no), acompañada de instantes de silencio y de ciertos momentos de incompreensión debido a registros lingüísticos diferentes. La narración personal de la experiencia de la enfermedad y del sufrimiento, del miedo, la vergüenza o la angustia no ocupan un lugar relevante dentro de esta interacción.

Esto lleva a reflexionar sobre el tipo de relación asimétrica que puede predominar dentro del espacio sanitario entre el médico y el paciente, en donde su “mirada médica” y su escucha se posan sobre un cuerpo-objeto de conocimiento. Un conocimiento basado principalmente en la enfermedad y en la recolección de toda una serie de datos y pruebas objetivas que, marcan a su vez la experiencia de la enfermedad (Chaiyachati, 2019). Encontramos detrás de todo ello, una racionalidad médica que actualmente y debido a lo tecnificada, hiper-especializada y burocrática que se encuentra la medicina, crea un espacio deshumanizado, frío e instrumental. Imperando así, un modelo biomédico hegemónico de corte positivista en donde el cara a cara con el paciente y la calidad de esta interacción no es uno de los aspectos más relevantes (Carrió, 2006).

Frente a esta forma de enfocar la praxis médica, en esta comunicación se abordará la propuesta del modelo de la medicina narrativa como alternativa metodológica de asistencia clínica. Un movimiento que surge en los años 90 del pasado siglo de la mano de distintos profesionales de la salud que se muestran críticos con este modo de enfocar la práctica médica. Siguiendo las aportaciones de su precursora Rita Charon en, “Narrative Medicine. A Model for Empathy, Reflection, Profession, and Trust” (2001), este modelo pone el énfasis en la importancia de las competencias narrativas de los profesionales sanitarios para promover una escucha más atenta del paciente y una ética de la atención y del

* natigarrido3@gmail.com

cuidado más humana y eficaz (Charon, 2001:1897). Se trata de una propuesta que apuesta por una escucha activa en la interacción entre el médico y el paciente para reconocer, interpretar y actuar sobre las historias de la enfermedad. Historias desde el punto de vista de los pacientes y diferentes a la del médico, que son reconocidas y correlacionadas para orientarlas hacia una relación que resulte beneficiosa, y que permita a la persona enferma dejar de sentirse como un mero objeto de conocimiento. A través de ello, el objetivo de este modelo de medicina narrativa es fomentar la dimensión intersubjetiva de este encuentro tendiendo puentes entre las distintas perspectivas que tanto el médico como el paciente tienen de la enfermedad, de sus síntomas y de la influencia en sus vidas (Charon, 2001:1899)

Desde este marco teórico, se explorará en esta comunicación el papel que asume en este modelo la narración de la enfermedad como “evidencia epistemológica” (Mariano, Rodríguez y Conde, 2013). La narración del paciente aporta al profesional sanitario información personal y contextualizada a la que no tendrían acceso de otra forma. Información que evita un diagnóstico desenfocado, que enriquece las praxis médica, que permiten una atención y una relación terapéutica más empática y eficaz y que ayuda a modificar determinadas pautas comunicativas asimétricas y estereotipadas dentro la práctica clínica.

Referencias bibliográficas

- [1] Carrió, S. (2006). “Aproximaciones a la medicina narrativa” en *Revista Hospital Italiano de Buenos Aires*, vol. 26, nº 1, marzo 2006, p. 14-19.
- [2] Charon, R. (2001). “Narrative Medicine. A Model for Empathy, Reflection, Profession, and Trust” en *JAMA*, vol. 286, nº 15, 17 de octubre de 2001, p. 1897-1902.
- [3] Chaiyachati, H., Shea, A., Asch, D., Liu, M., Bellini, L., Dine, J., Sternberg, A., Gitelman, Y., Yeager, A., Asch, J., y S. Desai (2019). “Assessment of Inpatient Time Allocation Among First-Year Internal Medicine Residents Using Time-Motion Observations”, en *JAMA Internal Medicine*, vol. 179, nº 6, junio 2019, p. 760- 767.
- [4] Guibert, H. (1992). *Cytomégalo virus. Journal d'hospitalisation*, Paris: Éditions du Seuil.
- [5] Guibert, H. (1991). *Le protocole compassionnel*, Paris: Éditions Gallimard. Citas de la edición castellana, *El protocolo compasivo*. Trad. Carlos Manzano, Barcelona: Tusquets editores, 1992.
- [6] Hunter, KM., Charon, R., y JL. Coulehan (1995). “The study of literature in medical education” en *Academic Medicine*, vol. 70, nº 9, 1995, p. 787-794.

- [7] MacLellan, F., y A. Hudson Jones (1996). “Why literature and medicine?”, en *The Lancet Journal*, vol. 348, nº 9020, julio de 1996, p. 109-111.
- [8] Mariano Juárez, L., Rodríguez Martín, B. y D. Conde Caballero (2013). “Cuidados basados en narrativas: redefiniendo la jerarquía de la evidencia” en *Index de Enfermería*, vol. 22(1-2), 2013, p. 55-59.
- [9] Tortolo, M. (2011). “Entrevista a la Dra. Rita Charon en Nueva York. Las historias que cuenta la medicina” en *IntraMed* (online), abril de 2011.

The Extended Evolutionary Synthesis and Scientific Incommensurability. Taking Kuhn Seriously

Juan Gefaell*
University of Vigo

Cristian Saborido**
UNED

Introduction

In the last few years, an increasing number of evolutionary biologists claim that the standard evolutionary theory, also called the Modern Evolutionary Synthesis (MES), ought to be reformed in order to account for several new biological findings. These reformist biologists have advocated for a so-called “Extended Evolutionary Synthesis” (EES; Pigliucci & Müller 2010; Laland et al. 2015). However, not all theorists agree that a profound reform of the classical evolutionary framework is really necessary (e.g., Futuyma 2017).

This debate between advocates and detractors of the rupture with the MES, has led philosophers to reflect on the conceptual issues that underlies the scientific discussion (e.g., Pigliucci and Finkelman 2014; Fábregas-Tejeda and Vergara-Silva 2018). Specifically, the philosophy of science attempts to determine whether EES really represents a radical departure from MES. The consequences of this debate in biology are particularly relevant from a philosophical point of view, insofar as it is a privileged case study for understanding scientific change.

Given the prominence philosopher Thomas Kuhn has had on the topic of scientific change, and that the terminology coined by him –e.g., terms such as “paradigm”, “normal science”, or “scientific revolution”– is frequently borrowed in philosophical accounts of the EES (e.g., Pigliucci 2012, 2017; Fábregas-Tejeda and Vergara-Silva 2018), we argue that a fine-grained Kuhnian analysis of the current EES vs. MES controversy is both interesting and necessary. In particular, due to frequent disputes around the degree of (dis)continuity between the EES and the MES, we argue that it is particularly useful to examine whether the two frameworks are truly incommensurable. To this end, it is important

* gefaell@uvigo.es

** cristian.saborido@fsof.uned.es

not only to know the details of each of the supposed “paradigms” competing in current evolutionary biology, but also to analyze what the Kuhnian notion of “incommensurability” really means. The moral of this paper is that taking Kuhn seriously for the current biological debate not only serves to recognize the inadequate uses of his ideas, but also to better understand the theoretical and practical implications of this discussion.

The MES and the EES

In order to determine if the MES and the EES are incommensurable, we first provide a brief characterization of these frameworks. On the one hand, the MES constitutes a synthesis between a mathematical theory of natural selection and a population-level Mendelism (Provine 1971), under which evolution is conceived as changes in the genetic composition of populations across generations (Futuyma 2005). On the other hand, the EES is composed by four biological principles (evo-devo, developmental plasticity, inclusive inheritance and niche construction), and two metascientific principles that exert direct influence in the evolutionary models and explanations of the EES: constructive development and reciprocal causation (Laland et al. 2015). We argue that, at bottom, these principles entail that organisms, and not genes, are the fundamental entities of evolution.

Incommensurability in Kuhn’s work

Then, we clarify what does incommensurability mean. We note that incommensurability was introduced by Thomas Kuhn in his classic *The Structure of Scientific Revolutions* (1962; SSR) in order to account for the relationship between two paradigms that are involved in a scientific revolution. According to Kuhn, would be no neutral basis from which to compare paradigms, making communication among scientists from different paradigms very difficult. We also emphasize that, in SSR, Kuhn posited that incommensurability would affect three domains of science: methodology, semantics, and observations. We briefly characterize these components of incommensurability as faithful as possible to Kuhn’s original discussion in SSR.

Besides, we show how, as several Kuhn scholars have noticed (e.g., Hoyningen-Huene 1990; Sankey 1993), the incommensurability concept changed after the SSR (Kuhn 1970, 1977, 2000). We summarize those changes in three points: (1) in his later works, the local nature of incommensurability was stressed out, (2) it became restricted to the semantic sphere, and (3) it became more precise. Indeed, we note that, in his later works, Kuhn provided two criteria with which to detect incommensurability. These two criteria are (Kuhn 2000: chap. 1): (a) in incommensurable theories, the most relevant categories are defined according to different criteria; and (b), the distribution of objects and properties within preexisting categories change.

Are the MES and the EES incommensurable?

Once we have analyzed both the MES and the EES, as well as the incommensurability concept, we tackle the issue if these theories are incommensurable. We claim that the best way to do it so is to ask two separate questions: (1) Are the MES and the EES incommensurable according to the SSR's version? And (2), are they so according to the late version?

In order to determine if the MES and the EES are incommensurable according to the SSR's version, we propose that it is necessary to check if those theories are methodologically, semantically and observationally incommensurable. We will defend that this is in fact the case, and we illustrate this point analyzing the explanatory styles, the definitions of "inheritance" and the ontological assumptions underlying organisms under each theory.

Regarding the late version of incommensurability, we check if the detected differences in concepts between these two theories stick to the criteria provided by Kuhn or not. We argue that, although there are significant differences between the concepts of each theory, these differences are not explained by the criteria provided by Kuhn. We illustrate this point with the concepts of "niche construction" and "objects of selection".

A conflict

As a conclusion, we argue that the two versions of the incommensurability concept render different results when applied to the MES vs. EES controversy. We show how this discrepancy has to do with the way in which each concept of incommensurability is defined: whereas the SSR's version is certainly vaguer and more comprehensive, making it easier to find instances in its favor, the late version is more restricted and more precise, making it harder to do so.

Our analysis poses the question of which version of incommensurability to choose, given the discrepancy of results. We defend that despite its shortcomings, the SSR's version is preferable. The reason is that, at bottom, the question about which version of incommensurability to choose is a question about preferences over epistemic values in the philosophy of science: we have, on the one hand, the SSR's version, which stands out for its comprehensiveness; and on the other hand, the version of Kuhn's later works, which does so for its precision. We argue that in those occasions in which the epistemic values of precision and comprehensiveness collide, we ought to choose comprehensiveness. As the recent pragmatist accounts have pointed out (e.g., Mitchell 2013), science is a complex thing; therefore, it is better to describe it vaguely but comprehensively, encompassing all aspects of scientific practice, than to do it in a precise manner but leaving aside key features that constitute its very nature, as the late version of incommensurability does.

Either way, choosing the incommensurability version of SSR, which implies that MES and EES are indeed incommensurable, seems to raise relevant issues. Among these issues, two stand out: first, if EES and MES are incommensurable, then are they not rationally comparable? And second, is evolutionary biology undergoing a Kuhnian scientific revolution? Our answers are “yes” and “no,” respectively. The analysis presented in this paper demonstrates that incommensurability does not imply rational incomparability, and that contemporary biology is not undergoing a Kuhnian revolution, even if one can agree that MES and EES are different enough to be considered incommensurable.

References

- [1] Fábregas-Tejeda, A., Vergara-Silva, F. (2018). “The emerging structure of the Extended Evolutionary Aynthesis: where does Evo-Devo fit in?” *Theor Biosci* 137:169-184.
- [2] Futuyma, DJ. (2005). *Evolution*. Sinauer Associates, Sunderland.
- [3] Futuyma, DJ. (2017). “Evolutionary biology today and the call for an extended synthesis”. *Interface Focus*, 7:20160145.
- [4] Hoyningen-Huene, P. (1990). “Kuhn’s conception of incommensurability”. *Stud Hist Philos Sci A*, 21:481-492.
- [5] Kuhn, TS. (1962). *The structure of scientific revolutions*. The University of Chicago Press, Chicago.
- [6] Kuhn, TS. (1970). “Reflections on my critics”. In: Lakatos, I., Musgrave, A. (eds). *Criticism and the growth of knowledge*. Cambridge University Press, Cambridge, pp. 231-278.
- [7] Kuhn, TS. (1977). *The essential tension: selected studies in scientific tradition and change*. The University of Chicago Press, Chicago.
- [8] Kuhn, TS. (2000). *The road since The Structure*. The University of Chicago Press, Chicago.
- [9] Laland, KN., Uller, T., Feldman, MW., Sterelny, K., Müller, GB., Moczek, A., Jablonka, E., Odling-Smee, J. (2015). “The extended evolutionary synthesis: its structure, assumptions and predictions”. *Proc Biol Sci*, 282:20151019.
- [10] Mitchell S (2013). *Unsimple truths: science, complexity, and policy*. University of Chicago Press, Chicago.
- [11] Pigliucci, M. (2012). *Biology’s last paradigm shift*. *Paradigmi*, 3:45-58.

- [12] Pigliucci, M. (2017). "Darwinism after the modern synthesis". In: Delisle RG (ed) *The darwinian tradition in context: research programs in evolutionary biology*. Springer International Publishing, New York, pp 89-103.
- [13] Pigliucci, M., Müller, GB. (eds). (2010). *Evolution -The extended synthesis*. The MIT Press, Cambridge.
- [14] Pigliucci, M., Finkelman, L. (2014). "The extended (evolutionary) synthesis debate: where science meets philosophy". *Bioscience*, 64:511-516.
- [15] Provine, WB (1971). *The origins of theoretical population genetics*. The University of Chicago Press, Chicago.
- [16] Sankey, H. (1993). "Kuhn's changing concept of incommensurability". *Br J Philos Sci*, 44:759-774.

Caracterización del concepto de dato en los ácidos nucleicos desde un enfoque diafórico

César González Herrerías*
Universidad de Oviedo

El uso de conceptos derivados de la teoría de la comunicación de Shannon (Shannon, 1948) es muy habitual en biología molecular. Términos como “información”, “mensaje”, “código” o “decodificación”, se utilizan con gran profusión en los trabajos de esta disciplina. No en vano, la idea de la transferencia de información desde el ADN a las proteínas constituye el núcleo conceptual del Dogma Central de la biología molecular contemporánea. Sin embargo, la justificación epistémica de su empleo continúa siendo objeto de permanente debate, fundamentado en la escasez de desarrollos teóricos y aplicaciones prácticas resultantes.

En cualquier caso, la cuestión de la utilización del vocabulario informacional en biología molecular no se reduce únicamente a la definición de un concepto de información característico y su posible justificación epistémica como herramienta para la descripción del proceso de síntesis proteica. La analogía informacional constituye un verdadero modelo general construido a partir de las posibles analogías particulares existentes entre los distintos elementos de un sistema de comunicación y las moléculas y estructuras biológicas que participan en los procesos de expresión génica. Por ello, resulta igualmente necesario justificar la utilización de cada uno de estos términos y sus interrelaciones, con el fin de validar, refutar o modificar el modelo en su conjunto. Para avanzar en la resolución de este problema se ha definido un marco teórico con el objeto de determinar las condiciones que habría de cumplir una analogía informacional general aplicada a la biología molecular que fuese consistente desde el punto de vista epistémico y coherente desde una óptica evolutiva.

En base al papel fundamental que desempeña el ARN en los procesos biológicos celulares y al lugar que ocupa en las teorías sobre el origen y la evolución temprana de la vida sobre la Tierra, se propone una redefinición de la analogía informacional que sitúa a la molécula de ARN como origen de la transmisión de información en los procesos celulares. Es decir, como análogo particular del elemento *fuentes de información* del proceso de comunicación descrito por Shannon. De este planteamiento se deriva otro consistente en que el genoma de un organismo no contendría información propiamente dicha, sino

* cesarherrerias@hotmail.com

únicamente datos, cuyo significado evolutivo, esto es, cuyo contenido informativo en términos de incremento de la aptitud biológica, solo emergería tras ser estos datos *interpretados* en virtud de los diferentes mecanismos de regulación de la expresión génica. Se adopta, por tanto, un enfoque semántico del concepto de información en biología molecular que entiende esta como “datos + significado biológico en términos evolutivos”. Para fundamentar esta idea nos basaremos en la Definición General de Información (GDI) propuesta por el filósofo de la información Luciano Floridi, así como en la visión de John Maynard Smith acerca de la selección natural como primer codificador de la secuencia de nucleótidos (Maynard Smith, 2000). De igual forma, consideraremos el trabajo del biólogo evolutivo Steven A. Frank relativo a las posibles equivalencias conceptuales y algebraicas existentes entre las ecuaciones que describen el cambio en las frecuencias génicas de una población y las ecuaciones de la teoría de Shannon. A su vez, la mencionada labor de *interpretación* será entendida como un mecanismo de procesamiento de los datos contenidos en el ADN, llevado a cabo por los circuitos lógicos celulares de regulación de la expresión génica.

La Definición General de Información establece tres condiciones que han de cumplirse para que algo pueda ser considerado como un ejemplo de información semántica: que ese algo conste de uno o más datos, que estos datos estén bien formados y que además sean significativos (Floridi, 2005). A su vez, Floridi propone la *Definición Diafórica de Dato*, en alusión al término griego *diaphora* “diferencia”, en la que se entiende el concepto de dato como un hecho putativo con respecto a alguna diferencia, discontinuidad o falta de uniformidad dentro de algún contexto (Floridi, 2019). Tal diferencia puede aplicarse a tres niveles ontológicos distintos: datos entendidos como *diaphora de re*, como *diaphora de signo* y como *diaphora de dicto*.

Los datos como *diaphora de re* se conciben como discontinuidades o faltas de uniformidad en el mundo real, físico. Se corresponderían con la raíz etimológica del término, es decir, con los *dedomena* de Euclides que fueron traducidos al latín como *data*, a saber, objetos geométricos, materiales al menos en parte, que son dados en primera instancia, permitiendo a su vez la determinación de otros objetos. Serían, en cierta medida, *la fuente de los datos*, esto es, datos puros o *protoepistémicos*, antes de ser captados o interpretados, a los que no se accede independientemente de cierto nivel de abstracción. Intentaremos justificar que estos pueden verse como datos relativos al paisaje adaptativo y bioquímico futuro en el que se desenvolverá un organismo, y más concretamente a la geometría tridimensional de proteínas, enzimas y ribozimas, así como de los sustratos con los que estas interactúan.

Por su parte, los datos como *diaphora de signo* se definen como una falta de uniformidad entre (la percepción de) al menos dos estados físicos diferentes (es el propio autor

quien pone entre paréntesis la expresión “la percepción de”). En este caso, lo definitorio no serían solamente los estados propiamente dichos, sino también la forma en que son percibidos o, desde nuestra perspectiva, *captados* o *codificados*. Siguiendo a Maynard Smith, identificaremos al proceso de selección natural como el sistema que hace posible tal codificación en el genoma, y a la diferencia, discontinuidad o falta de uniformidad existente entre los pares de características “base púrica/base pirimidínica”, “grupo amino/grupo ceto” y “tres puentes de hidrógeno/dos puentes de hidrógeno” de las bases nitrogenadas, como los pares de estados físicos, o *forma de codificación*, a los que se hace mención en la definición.

Por último, el concepto de dato como *diaphora de dicto* se define como una diferencia o falta de uniformidad entre dos símbolos (Floridi, 2019). Se alude aquí al carácter simbólico de los datos, entendiendo estos, por tanto, como una forma de representación de otras entidades. Con el fin de aplicar esta noción a una secuencia de bases de ADN o ARN, distinguiremos dos categorías de símbolos en el repertorio propuesto por el Comité de Nomenclatura de la IUPAC-IUBMB¹ para los ácidos nucleicos. La primera de ellas estaría formada por aquellos símbolos que poseen un referente bioquímico real. Esto es, o bien la propia base nitrogenada que se encuentra *efectivamente* en una posición de la cadena, o bien una característica química relevante de esta. La segunda categoría la formarían el resto de los símbolos que designan una base no especificada totalmente, dentro de un conjunto de tres o cuatro posibles bases. Se trata en este caso de una recodificación que tiene como objeto representar, de forma comprimida en un único símbolo, las distintas posibilidades que se pueden dar en una posición en concreto de una secuencia de ADN o ARN. En base a esta distinción intentaremos justificar que las posibles *diaphora de dicto* derivadas de la presencia de una base en una posición de la cadena pueden reducirse únicamente a tres: “R⊕Y”, “M⊕K” y “S⊕W”. Es decir, aquellas que expresan la diferencia o falta de continuidad existente entre dos símbolos que poseen como referente una característica química real de las bases nitrogenadas, denotadas mediante una relación de disyunción exclusiva.

Finalmente, propondremos que el enfoque diafórico puede verse como un proceso recursivo de la definición de información semántica, en el que se aplica la misma función a un problema de idéntica naturaleza pero menor complejidad, partiendo de un problema de base irreductible que se correspondería con los datos como *diaphora de re*. Así mismo esbozaremos las interrelaciones entre este enfoque y los planteamientos de la biosemiótica.

¹ Siglas en inglés de la International Union of Pure and Applied Chemistry [Unión Internacional de Química Pura y Aplicada] y la International Union of Biochemistry and Molecular Biology [Unión Internacional de Bioquímica y Biología Molecular].

Bibliografía

- [1] Floridi, Luciano (2005). “Is Information Meaningful Data?” *Philosophy and Phenomenological Research*, 70 (2): 351-370.
- [2] Floridi, Luciano (2019). “Semantic Conceptions of Information”, *The Stanford Encyclopedia of Philosophy*, recuperado el 7 de mayo de 2021 de <https://plato.stanford.edu/archives/win2019/entries/information-semantic>.
- [3] Frank, Steven (2012). “Natural Selection. V. How to read the fundamental equations of evolutionary change in terms of information theory”, *Journal of evolutionary biology*, 25: 2377-96.
- [4] Maynard Smith, John (2000). “The Concept of Information in Biology”, *Philosophy of Science*, 2: 177-194.
- [5] Shannon, Claude Elwood (1948). “A Mathematical Theory of Communication”, *Bell System Technical Journal*, 27 (4): 623-666.

Análisis de la condición epistémica de la responsabilidad moral desde “tres enfoques” para el contexto tecnológico

Joshua Alexander González Martín*
Universidad de Salamanca

Desde la obra de Aristóteles *Ética a Nicómaco*, la tradición entre los filósofos morales ha sido identificar dos condiciones individualmente necesarias y conjuntamente suficientes para que alguien sea moralmente responsable de algo. Primero, hay una condición de voluntad o control; en segundo lugar, existe una condición epistémica. Recientemente, la condición epistémica se ha convertido en el foco de un mayor escrutinio.

El objeto epistémico de estudio ha variado con los problemas éticos y tecnológicos con los que actualmente lidiamos. Si para algunos la condición epistémica habla de que la responsabilidad moral requiere de conocimiento de lo que se está haciendo, para otros la condición epistémica habla de un conocimiento moral, de la posibilidad de tener creencias morales verdaderas o falsas. A la primera consideración la he llamado enfoque psicológico y, a la segunda, enfoque ético. En cambio, la condición epistémica también puede estar relacionada, por ejemplo, con el conocimiento de los riesgos y resultados por la acción tecnológica. A este enfoque lo he llamado, propiamente, enfoque epistémico. El enfoque epistémico estudiará este tipo de cuestiones: ¿Cómo podemos atribuir una responsabilidad moral cuando los efectos de las acciones son (o han sido) difíciles de conocer o simplemente desconocidos?

Analizo brevemente las condiciones epistémicas de la responsabilidad moral en estos tres enfoques y adelanto la tesis de que, en contextos tecnológicos, sólo es excusable la ignorancia cuando no se tiene ningún tipo de posibilidad de acceso a evidencias por el que no actuar como se ha actuado.

Según el enfoque que llamamos psicológico, la responsabilidad moral requiere de voluntad y esta está estrechamente relacionada con el control y el conocimiento. La responsabilidad así entendida forma parte de un sistema constituido por agentes que actúan mediante una relación de voluntad y conocimiento con la acción que cometen (y, a veces, con las posibles alternativas a esta), requisitos para que los agentes sean responsables de sus acciones y por los que las acciones serán por completo imputables a los agentes. Por lo tanto, acción que no sea realizada conscientemente no puede ser intencional; acción que no es intencional no puede ser libre; y acción que no es libre no puede ser imputada

* u159840@usal.es

a un agente, el cual tampoco podrá ser responsable de ella si su acción no contaba con alternativas.

En este sentido, ya desde Aristóteles se propone que los seres humanos realizan “acciones voluntarias” y, por tanto, son responsables de lo que hacen (Eth. Nic. III, 2, 1111b). Así, un agente es elogiado o acusado (legítimamente) sólo por cosas que son voluntarias. Esta tesis psicológica sobre la responsabilidad moral sigue siendo un asunto actual y se sigue aceptando que la responsabilidad moral requiere la satisfacción de alguna condición mental, aunque se ha discutido mucho acerca de cuál podría ser esta condición mental. En la línea tradicional, algunos han afirmado que el control requiere conciencia y han mostrado que ambas condiciones se contienen (Sher, 2009: 146). Con respecto a la condición de libertad en particular, algunos han afirmado que la responsabilidad moral no requiere control (Smith, 2005). Y en cuanto a la condición epistémica, algunos han afirmado que la responsabilidad moral no requiere conciencia (Sher 2009).

Sin embargo, está muy extendida la idea de que, para ser moralmente responsable de algo, una persona debe estar en un cierto estado mental y ser capaz de elegir la acción correcta mediante algún tipo de conocimiento. El criterio de la historia causal ha sido, hasta cierto punto, satisfactorio para explicar en qué sentido una persona es responsable de su acción (Zimmerman, 1997). El caso de actuar en estado de alucinación voluntaria es un ejemplo común en este contexto. Piense en un agente que, conscientemente y por voluntad propia, ingiere algún tipo de estupefaciente que le hace sufrir alucinaciones y perder el control. El agente está tan intoxicado que empieza a ver monstruos donde hay personas y atacantes donde hay personas acercándose preocupadas. Este estado provoca que el agente se asuste y comience a lanzar golpes para defenderse de sus atacantes imaginarios, infligiendo daño a esas personas. El agente estaba tan drogado en el momento de la golpiza que no era capaz de formar voliciones que le permitieran actuar de otra manera. Sin embargo, consideramos que el agente es responsable de los daños porque anteriormente decidió conscientemente y de forma voluntaria (con posibilidad de alternativa) ingerir los estupefacientes.

Cuando actuamos, a menudo ignoramos algún hecho relevante. Sin embargo, el enfoque que llamamos ético difiere del primero en la pregunta: ¿saben los agentes que su conducta es moralmente incorrecta? En el ejemplo, el agente conoce porque consume estupefacientes voluntariamente y es consciente de ese consumo. En cambio, el enfoque ético habla de que los agentes conocen (o ignoran) que su conducta es moralmente incorrecta (Wieland, 2017). El agente sabe que consumir estupefacientes tiene ciertas consecuencias como la que tuvo y sabe que podría haber actuado de otra manera, pero desconoce que sus acciones y las consecuencias de sus acciones son moralmente incorrectas (Sher, 2017).

En esta línea, dados los recursos intelectuales y culturales de los que disponemos en un contexto concreto, como los agentes que operan en un nuevo sistema de producción

con incidencias ambientales nuevas, será necesario un genio moral para darse cuenta de lo erróneo del nuevo sistema (Rosen, 2003). En cambio, en las situaciones en las que se tiene (o se tiene la obligación de acceso al) conocimiento moral, se tendrá (o se deberá tener) una creencia verdadera justificada. En cuanto al deber de investigación ética, un agente puede ser culpable de diferentes cosas, incluida su omisión para informarse a sí mismo. Así, una empresa de biotecnología puede desarrollar “semillas terminator”. Ahora bien, antes de comercializarlas, ¿la empresa tendría el deber de estudiar las evidencias morales, para así actuar de la forma más correcta posible? ¿Sabían que podían hacerlo y que no hacerlo suponía un riesgo injustificado de sufrir algún daño? Si hay evidencia en el mundo y puede llegar a tenerla a través de cualquier esfuerzo que pueda hacer, entonces parece que está en el poder de la empresa llegar a tener una creencia verdadera justificada. El criterio moral, en este sentido, es un tipo de conocimiento ético que puede aportar fundamento a algunas respuestas.

El criterio del conocimiento moral puede trasladarse a otros tipos de conocimientos. Es a esto a lo que llamamos propiamente enfoque epistémico. Las preguntas que se plantean desde este enfoque son del tipo: ¿Los agentes deberían haberse informado sobre algún asunto? ¿Qué grado de ignorancia es necesario para que un agente no sea moralmente responsable de algo? Piense en uno de los mundos posibles, donde la empresa CloudMinds ha desplegado una versión altamente sofisticada de máquinas inteligentes conectadas a la nube a través del sistema Human Augmented Robotics Intelligence. Estas máquinas aprenden a partir de la demostración humana. Sin embargo, la conducta humana no siempre es la correcta y corremos el riesgo de que esta tecnología proporcione una agenda de acciones incorrectas a los sistemas tecnológicos conectados.

Por supuesto, la acción incorrecta de esta tecnología no era la intención empresarial; al contrario, la intención es proporcionar inteligencia artificial fiable y seguridad de redes. De hecho, la empresa ignoraba el riesgo de acciones incorrectas y está devastada por el resultado. En estas condiciones, la mayoría diría que CloudMinds no tiene la culpa de los daños causados, a menos que tenga la culpa de su ignorancia (Zimmerman, 1997). La ignorancia es no conocer las evidencias. Para el caso de CloudMinds, no sería preciso decir que cree en una verdad injustificadamente, pues suponemos que la empresa no actuó con la creencia injustificada del riesgo y, haciendo caso omiso, procedió de todos modos como lo hizo. La culpa, en este caso, no sería excusada. En cambio, cuando nos preguntamos por la ignorancia culpable, nos estamos preguntando si se puede culpar a los agentes por los resultados perniciosos de los actos que realizan a razón de tener creencias falsas sobre algún asunto.

Como conclusión, lo dicho hasta y la posibilidad de extender las sutilezas de cada particular, permiten una nueva integridad conceptual para abordar el criterio epistémico de la responsabilidad en la tecnociencia contemporánea. El valor de la responsabilidad,

en este sentido, podrá ser ajustado para su atribución tanto a cualquier ser humano técnicamente consciente como, particularmente, a los tecnólogos, los políticos, las empresas o los militares, que tienen el poder del acceso al conocimiento.

Referencias bibliográficas

- [1] Aristóteles (2019). *Ética a Nicómaco*, Madrid: Gredos.
- [2] Rosen, G. (2003). “IV– Culpability and Ignorance”. *Proceedings of the Aristotelian Society*, 103(1), 61-84.
- [3] Sher, G. (2009). *Who knew?: Responsibility without awareness*. Oxford University Press.
- [4] Sher, G. (2017). “Blame and moral ignorance”. En P. Robichaud & JW Wieland (ed.), *Responsibility: The Epistemic Condition*, 101-116.
- [5] Smith, A. M. (2005). “Responsibility for attitudes: Activity and passivity in mental life”. *Ethics*, 115(2), 236-271.
- [6] Wieland, J. W. (2017). “The epistemic condition”. En P. Robichaud & JW Wieland (ed.), *Responsibility: The Epistemic Condition*, 1-28.
- [7] Zimmerman, M. J. (1997). “Moral responsibility and ignorance”. *Ethics*, 107(3), 410-426.

Not So Distinctively Mathematical Explanations: Microphysics matters

Aditya Jha*

University of Canterbury, New Zealand

A number of modally robust distinctively mathematical explanations (DMEs) have been proposed in the literature which appeal primarily to mathematical necessities rather than contingent causal laws. This paper has two broad goals:

- (a) show that the two putative case studies of non-causal DMEs of complex systems – the double pendulum (Lange 2016) and antipodal weather patterns (Colyvan, 2001; Lange, 2016) – fail as DMEs, due to failure of modal robustness, on introducing certain physical and mathematical counterfactuals that make use of the microphysical richness of such systems, and
- (b) highlight the competing values in such DMEs – namely, the distinctiveness desideratum, the modal desideratum, and the directionality desideratum (Povich, 2019) – by introducing the ‘epistemic desideratum’ which illuminates the force of the modal desideratum. This is motivated by the obvious, but largely neglected, question in the literature: why are DMEs even applicable to target systems? (Molinini, 2021).

Part I of the paper gives an overview of the different kinds of DMEs of double pendulum systems, which this paper extends to cover any n-tuple pendulum, and then shows why each of them fails on account of their failure to satisfy these competing values. Part II uses various case studies to suggest that the microphysical richness of atmospheric systems block the assumptions required to exploit the mathematical theorems that make the DME on antipodal weather patterns true.

Part – I: Pendulum Systems

The causal explanation of the number of equilibrium positions of a double pendulum uses partial derivatives of a potential energy function with respect to the bob’s inclination. Lange (2016) proposes that there is also a non-causal DME which appeals primarily to topological facts about the pendulum system – its configuration space, which constrains its number of equilibrium positions – and thus unifies the explanation of such a constraint for any double pendulum, something that cannot be promised with a causal explanation.

* aditya.jha@pg.canterbury.ac.nz

The paper refutes not only Lange's (2016) DME for double pendulums (D1 below) but also various other topological DMEs of pendulum systems (which this paper introduces). To briefly introduce the DMEs, they are:

- (D1) There are four or more equilibrium positions of any double pendulum because its configuration space is a torus which has an invariant Euler characteristic (Lange 2016, pp. 27-31).
- (D2) There are four or more equilibrium positions of any double pendulum because for any n-tuple pendulum the lower bound on the number of equilibria with k stable directions is equal to the kth Betti number of its configuration space (n-dimensional torus), where a stable direction implies a pendulum rod pointing downward.
- (D3) The number of equilibrium positions for any n-tuple pendulum is at least 2^n because of binomial induction.
- (D4) The number of equilibrium positions for any n-tuple pendulum is at least 2^n because each pendulum bob can take exactly two positions under the zero line of force acting on the bob.

DISCUSSION on (D1-D4): The paper first discusses (D1) and shows that it is applicable non-trivially only to a simple or a double pendulum and gives trivial results for the constraint on the number of equilibrium positions when the strategy backing (D1) is extended to any other kind of pendulum systems such as triple or higher pendulum systems. Then the alternative strategy (D2) is proposed which is applicable to any n-tuple pendulum system non-trivially in predicting an exact lower bound for the number of equilibrium positions. Despite the modal strength promised by (D1) and (D2), the paper shows that they are only applicable to potential energy functions that are Morse functions (with non-degenerate Hessian matrix). Both the explanations fail for non-Morse functions (with degenerate Hessian matrix), such as when non-linear length perturbations are introduced in the pendulum rod say due to microphysical non-stiffness: for such functions, it is not possible to assign a general invariant Euler characteristic and, thus, the topological explanation fails to work. In terms of competing values, it is shown that both (D1) and (D2) fare well on the epistemic desideratum but fail the modal desideratum since their robustness is not guaranteed under perturbations.

The explanations (D3) and (D4) fare better on the modal desideratum since they are applicable under both Morse functions and non-Morse potential energy functions. But (D3) is shown to have poorer epistemic hooking than (D1), (D3), and (D4) because it has poor directionality and does not address the reason why binomial induction should explain anything about a pendulum system. (D4) fares better on the epistemic desideratum since the number of equilibrium positions or intersection points along the zero line of force can be explained by the geometry of the pendulum system. But this explanation is

denied as a DME as well because it is not clear why it should not be construed as a merely geometrical explanation of geometric facts, i.e., the explanation fails on the ‘distinctiveness’ desideratum since the mathematics does not seem to play any role in the explanation rather than indexing physical facts.

DMEs promising greater modal strength (and wider applicability) come with poor epistemic hooking, and DMEs with lesser modal strength (and narrowed applicability) come with the right kind of epistemic hooking. Thus, there are no DMEs of any n-tuple pendulum systems since they collapse on grounds of modality, directionality, or distinctiveness.

Part – II: Antipodal Weather Patterns

Colyvan (2001) argues that there are at least two antipodal points on earth that have the same temperature and pressure at a given time. He argues that such events are hard to explain in a unified way via causal mechanisms, but can be readily unified and explained by a theorem in algebraic topology – the Borsuk-Ulam theorem with two ‘minor’ assumptions, that the earth is a sphere and that temperature is a continuous function. Lange (2016) argues that this is a non-causal DME and analogous to the topological explanation of the double pendulum case since the topological theorem holds true regardless of the constitution of physical laws. Alternatively construed, the claim is that no micro-physical or causal detail (other than violating framework laws such as Newton’s second law or the laws of mass/energy conservation) can render this explanation incorrect.

The paper refutes this claim and argues that micro-physical details of temperature distribution and heat transfer matter to the explanation. Assuming temperature as a continuous function is a non-trivial assumption failing which the explanation clearly fails since the Borsuk-Ulam theorem works only for continuous functions.

The paper discusses the case of atmospheric heat transfer in fluids, such as vapor, and argues that the assumption of continuous functions is based on the continuum hypothesis in fluid mechanics. This hypothesis is applicable only when the mean free path (M) of the fluid molecules is considerably smaller than the length (L) of the Representative Volume Element (RVE), i.e. the Knudsen number ($K=M/L$) is very low such as $K < 0.1$. For higher Knudsen numbers, the continuum hypothesis is invalidated because of the introduction of slip and transition effects in addition to free molecular flows in the fluid.

The paper discusses two such sub-cases – rarefied gas dynamics and microfluidic flows – where K is relatively higher and discrete analysis of the distribution of molecular heat transfer is preferred because temperature changes are discrete in RVEs that are comparable to the length of M . In addition, studies of significant temperature discontinuities observed at the evaporative water interface are discussed (Fang & Ward 1999) and it is

shown how the assumption of temperature continuity (Dunn et al. 2008) has actually prevented a better understanding of the evaporation dynamics.

The paper thus notes that a blanket DME that ignores the microphysics of heat transfer and temperature discontinuity fails on epistemic grounds and thus has no modal force whatsoever. “The problem of the physical infinitesimal [concerning microphysics]” (Wilson 2017, Ch. 3) makes itself significant and underlines the importance of having an epistemic grasp on such explanations. The competing values of having an epistemic grasp vis-à-vis modal robustness thus deserve greater attention in the literature than they have previously received.

References

- [1] Colyvan, M. (2001). *The Indispensability of Mathematics*. Oxford: Oxford University Press, pp. 49-50.
- [2] Lange, M., (2016). *Because Without Cause: Non-Causal Explanations in Science and Mathematics*. Oxford University Press: USA.
- [3] Dunn, G. J.; Wilson, S. K.; Duffy, B. R.; David, S.; Sefiane, K. (2008). “A mathematical model for the evaporation of a thin sessile liquid droplet: comparison between experiment and theory”. *Colloids Surf., A*, 323, 50-55, DOI: 10.1016/j.colsurfa.2007.09.031
- [4] Fang, G.; Ward, C. A. (1999). “Temperature measured close to the interface of an evaporating liquid”. *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.* , 59, 417– 428, DOI: 10.1103/physreve.59.417
- [5] Molinini, D. (2021). “The Unreasonable Effectiveness of Physics in Mathematics”. *British Journal for the Philosophy of Science*. <https://doi.org/10.1086/715104>
- [6] Povich, M. (2019). “The Narrow Ontic Counterfactual Account of Distinctively Mathematical Explanation”. *British Journal for the Philosophy of Science*:axz008.
- [7] Wilson, M. (2017). *Physics Avoidance: Essays in Conceptual Strategy*. Oxford: Oxford University Press.

Descartes on Mathesis Universalis and on Method

Ladislav Kvasz*

Czech Academy of Sciences, Institute of Philosophy

John Schuster in his book *Descartes Agonistes* (Schuster 2013) offered a detailed and persuasive interpretation of the development of Descartes' thought from his early encounter with Beeckman until the writing of *Le Monde and the Discourse on Method*. The book is thus an important contribution to our understanding of the intellectual development of the early Descartes.

Schuster characterizes Descartes' method from his Discourse as method talk. He characterizes Descartes' description of his method as an ideologically charged form of discourse that by means of particular literary effects creates an illusion that it, for reasons rooted in the very structure of that discourse, cannot fulfill. I consider the notion of method talk a rather illuminating concept. When applied to the works of figures like Lakatos or Popper such an analysis can explain several features of the methodological debates in the 1960-ties. Thus in general I consider Schuster's approach to methodology as interesting and worth of further development.

Nevertheless, its application to Descartes seems to me misleading. Descartes was one of the greatest mathematicians of all times. He was of the same rank as Newton, Euler, Riemann, or Poincaré. The reason why many historians do not see him that way is that his mathematical innovations (in contrast to those of Euler, Riemann, or Poincaré) became the basis of secondary mathematics education, and so we are used to them to such a degree, that they appear to us as almost trivial. But Descartes' idea to visualize a polynomial is as radical and as revolutionary as Euler's introduction of a complex variable into the calculus, Riemann's move to n -dimensional manifolds in geometry or Poincaré's invention of algebraic topology. Therefore I believe we stop for a while before we dismiss Descartes' views on method as a method talk comparable to Lakatos' or Popper's discussions of method.

The aim of my paper is to invite the reader to such a stop and offer an alternative reconstruction of the mathematical meaning of Descartes' method. I agree with Schuster in most of the historical details and textual evidence he presents. Nevertheless, I will try to clarify the mathematical meaning of the particular stages of the development of Descartes' views on method. Schuster basically reconstructs Descartes' development towards the method talk in three stages. I accept these stages, but I will argue that these stages can

* ladislavkvasz@gmail.com

be characterized in a more technical way, which will reveal, that what Descartes offered in the particular stages was definitely not a form of discourse built on literary effects as Schuster reconstructed them.

To refute Schuster I will try to show that Descartes' position during each of these stages (there were in fact four of them) makes from a thoroughly mathematical point of view a very good sense. Nevertheless, in order to understand this mathematical meaning, one has to put Descartes' views not into the context of his contemporary intellectual situation in Paris centered around scepticism, but we have to put him on the landscape of mathematics proper, thus positioning his efforts into context of those of Euler, Riemann, and Poincaré. (I am here, of course, systematically suppressing the name of Newton, because there was a tension between the style of Descartes' mathematics and of that of Newton, but if the reader is able to overcome this tension, then of course, the first name on the list, standing before Euler, should be that of Newton).

The first stage of the development of Descartes' method Schuster identifies by his early involvement with *mathesis universalis*. Schuster characterizes this stage by Descartes' work with proportional compasses. Nevertheless, he sees in the compass only an idiosyncratic tool with a limited scope of applicability and so dismisses Descartes' universal mathematics as a literary genre borrowed from late antiquity. But if we put the proportional compass into a mathematical context we find that:

1. Antiquity knew only a very limited number of curves, actually only about a dozen (as the Conchoid or the Cissoïd). So when Descartes introduced an infinite system of different curves, it was a great advancement of geometry. That we do not see it this way is because later he introduced something even better, namely analytic geometry. But more importantly:

2. Why is the compass a tool of universal mathematics? Schuster is not posing this question, which is the key to understanding of this first stage. If we consider the curves of the ancients, the compositional synthesis of them is "arithmetical", e.g. in the case of the Cissoïd, we rotate a line and make sure that a segment is equal to another one. Thus the curve has a synthesis of the form $1+1+1+\dots$. In contrast to this the magnitudes in the proportional compass are in proportion. So the compositional synthesis is "algebraic", passing from x_1 to $x_2, x_3, x_3 \dots$ instead of the arithmetical passing from 1 to 2, 3, 4.... This new kind of unity of elements forming the object is crucial. Descartes understood it. Later he made this kind of synthesis obsolete by introducing coordinate constructions.

3. The compass radically enriches the expressive power of geometry. Synthetic geometry operates in fields that are quadratic extensions of the field of rational numbers. When Schuster characterizes Descartes' proportional compass by its power to solve cubic equations, he presents it as a limited innovation. But the compass can introduce also roots

of higher degrees, not only the third roots, thus it can capture all radical extensions. So in a sense it can do everything we can do in elementary algebra.

Thus Descartes had good reason to consider his proportional compass a basis of a universal mathematics that could really work. In a similar vein I will argue that the three remaining stages of Descartes' road towards his method had a working mathematical core. Schuster considers Descartes' methodological ideas only as method talk. In the paper I will argue that the mathematical core, that can be identified at each stage of Descartes' development, turn Schuster's use of the concept of method talk into an ideologically charged form of discourse that by means of particular literary effects creates an illusion (namely that Descartes uses method talk) that it cannot justify.

Las metáforas científicas y nuestra concepción de la Tierra

Henar Lanza*
Universidad del Norte

En su *Novum Organum*, Francis Bacon recurrió a la metáfora para transmitir su concepción de la nueva ciencia. En la primera parte de la obra (18, 21, 67 y 112) encontramos metáforas que comparten la idea de penetrar (penetretur) las entrañas de la Naturaleza, desflorarla.

Precisamente, es el propio Bacon el que en la misma obra nos advierte de que a pesar de que solemos creer que es la razón la que manda en las palabras, lo que ocurre es lo contrario, que son las palabras las que “ejercen a menudo una influencia poderosa sobre la inteligencia, lo que hace la filosofía y la ciencia sofisticadas y ociosas.” (I, 59).

El poder de estas metáforas determinó la concepción de la ciencia moderna, lo que determinó la relación de la especie humana con la Tierra: la correspondencia entre scientia y potentia se tradujo en una relación de dominio sobre la naturaleza. Pero no solo sobre la naturaleza, sino también sobre el ser humano, y así lo denunciaron Horkheimer y Adorno en *Dialéctica de la Ilustración* (1944), para quienes la Ilustración se había convertido en un mito.

En esta línea crítica, pero desde el ecofeminismo, Carolyn Merchant, química e historiadora y filósofa de la ciencia, analizó en *The Death of Nature. Women, Ecology, and the Scientific Revolution* (1980) las metáforas usadas por Bacon para describir el método científico, influidas, según ella, por el hecho de que Bacon asistió a los juicios a los que se sometía a las mujeres acusadas de brujería. En su análisis, la revolución científica del siglo XVII trajo consigo un cambio de paradigma (Kuhn, 1962): la cosmología organicista premoderna fue sustituida por una mecanicista. Esta conceptualización de la Tierra no como un organismo vivo, sino como una máquina, tuvo como consecuencia la dominación de la naturaleza y de las mujeres en nombre del progreso, la industria y el comercio. La imagen de Madre Tierra que había ejercido de restricción moral fue minada por la actividad comercial de la Edad Moderna y, una vez dejada atrás la imagen de madre nutricia, cayeron los obstáculos para el extractivismo, la minería y la deforestación.

Ya entrada la década de los 80, el empirismo feminista desarrollado por Sandra Harding, planteó en *The Science Question un Feminism* (1986) que dichas metáforas habían pretendido hacer atractivas moral y políticamente las nuevas concepciones de la

* lanzam@uninorte.edu.co

naturaleza y de la investigación que exigía el método inductivo con base en la recolección de experiencias. Sin embargo, la cuestión es qué relevancia tienen estos escritos para la práctica científica contemporánea y qué justifica que debamos considerar dichas metáforas como componentes fundamentales de las explicaciones científicas, más cuando el dominio de la naturaleza se ha traducido en un acceso desigual a los recursos naturales con el fin de la dominación social. Si la ciencia no es una herramienta al servicio de la emancipación universal, ¿por qué se presenta el nacimiento de la ciencia moderna como un momento progresista en la historia humana?

Teniendo en cuenta los análisis de la teoría crítica, Merchant y el empirismo feminista, la propuesta de esta investigación es que es necesario:

- dejar de ignorar las anomalías a las que conduce el actual paradigma de dominio y explotación;
- sustituir el paradigma mecanicista y las metáforas baconianas;
- tener en cuenta el análisis de la metáfora hecho ya por Aristóteles en su *Poética*, Blumenberg (2003), Kristeva (1984), Yurkievich (1979), Lizcano (2006) y Della Dora, V. (2021);
- desarrollar ontologías horizontales y relacionales como hicieron Arne Naess (1973), Bruno Latour (2011);
- evitar cualquier esencialismo –de género, raza o especie– y adoptar perspectivas postgenéricas e interespecies, como Donna Haraway (1991 y 2016);
- reavivar la relación humano-humus-humildad reivindicada por biólogas como Rachel Carson en *Silent Spring* (1962), Lynn Margulis en *Microcosmos* (1986) y Donna Haraway en *Staying with The problem* (2016), quien se declara no posthumanista, sino compostista;
- subrayar la importancia de que haya sido la biología la ciencia que haya rescatado la imagen antigua del hombre nacido del polvo, de la tierra o del barro que encontramos, por ejemplo, pero no solo, en el Génesis, en el mito de los metales de la República de Platón, en el mito del gigante Anteo;
- pensar el humus y el compost con los atributos de la chóra del Timeo de Platón, una matriz indeterminada y bastarda en la que ya están todas las formas de vida y que, en tanto tal, obliga a una replantear los usos del suelo, las aguas y los pesticidas.

Referencias

- [1] Adorno, Th. y Horkheimer, M. (1997). *Dialéctica de la Ilustración*. J. José Sánchez (trad). Madrid: Trotta.
- [2] Aristóteles (1997). *Poética*. Valentín García Yebra. Madrid: Gredos.
- [3] Bacon, F. (1984). *Novum organum*. Cristóbal Litrán (trad). Madrid: Sarpe.

- [4] Blumenberg, H. (2003). *Paradigma para una metaforología*. Jorge Pérez de Tudela (trad.). Madrid: Trotta.
- [5] Carson, R. (2013). *Primavera silenciosa*. Joandomènec Ros (trad.) Barcelona: Crítica.
- [6] Della Dora, V. (2021). *The Mantle of the Earth: Genealogies of a Geographical Metaphor*. Chicago and London: University of Chicago Press
- [7] Harding, S. (1986). *The science question in feminism*. Ithaca, NY: Cornell University Press.
- [8] Haraway, D. (1991). "A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century." *Simians, Cyborgs and Women: The Reinvention of Nature*, Nueva York: Routledge, 1991. pp. 149-181.
- [9] Haraway, D. J. (2016). *Staying with the Trouble: Making Kin in the Chthulucene*. Durham: Duke University.
- [10] Kristeva, J. (1984). *Revolution in poetic language*. New York: Columbia University Press.
- [11] Kuhn, T. S. (1962). *The structure of Scientific Revolutions*. Chicago - London: Chicago Press.
- [12] Latour, B. (2011). *Esperando a Gaia. Componer el mundo común mediante las artes y la política*. <http://www.bruno-latour.fr/sites/default/files/downloads/124-GAIA-SPE-AP-SPANISHpdf.pdf>
- [13] Lizcano, E. (2006). *Metáforas que nos piensan*. Madrid: Traficantes de sueños.
- [14] Margulis, L. y Sagan, D. (1986). *Microcosmos*. Mercé Piqueras. (trad.) Barcelona: Tusquets. [1985]
- [15] Merchant, C. (1980). *The Death of Nature. Women, Ecology and the Scientific Revolution*. New York: Harper Collins.
- [16] Merchant, C. (2008). "Secrets of Nature: The Bacon Debates Revisited". *Journal of the History of Ideas*, 69(1), 147-162. <http://www.jstor.org/stable/30139672>
- [17] Merchant, C. (2013). "Francis Bacon and the 'vexations of art': Experimentation as interventio". *British Journal for the History of Science*, 46(4), 551-599. <http://ezproxy.uninorte.edu.co:2084/10.1017/S0007087411000665>
- [18] Naess, A. (1973). "The shallow and the deep, long-range ecology movement. A summary". *Inquiry*. 16:1-4, 95-100. DOI: 10.1080/00201747308601682

- [19] Platón (1986), *Diálogos. IV. República*. C. Eggers Lan (trad.) Madrid: Gredos.
- [20] Platón (1992). *Diálogos. VI. Timeo. Filebo. Critias*. M.Á. Durán y F. Lisi (trad.) Madrid: Gredos.
- [21] Yurkievich, S. (1979). “Altazor. La metáfora deseante”. *Revista Iberoamericana*, Vol. XLV, Núm. 106-107, enero-Junio.
<http://revista-iberoamericana.pitt.edu/ojs/index.php/Iberoamericana/article/view/3361/3540>

Conocimiento científico y toma de decisiones: (in)certidumbre de la evidencia en la ciencia reguladora

Roberto López Mas*
Universidad de las Islas Baleares

Introducción

¿Qué evidencia es suficiente o necesaria para considerar probada una hipótesis? Esta es una cuestión de amplio debate en los últimos años, sobre todo por la importancia de las posibles respuestas para disciplinas como la medicina o la ciencia reguladora. Agencias como la EFSA han determinado que una asociación hipotética se define como una relación causal cuando existe evidencia directa generada por estudios de intervención en humanos. Por otro lado, autores críticos con esta propuesta defienden que las relaciones causales se pueden establecer mediante diferentes tipos de evidencia procedente de estudios con diseños distintos. En este trabajo se analiza, en primer lugar, la controversia metodológica en torno al papel de los RCT para la evaluación de hipótesis. Además, se estudia el debate concreto en la ciencia reguladora, donde no solo está en juego el conocimiento científico, sino también valores no epistémicos como la salud pública.

Metodología y evidencia científica

En las ciencias de la salud, el ámbito de la nutrición y la ciencia reguladora, es habitual el uso de jerarquías evidenciales que ordenan los distintos tipos de métodos en función de su capacidad para generar datos con un mayor o menor nivel de fiabilidad. El establecimiento de estas jerarquías se justifica sobre la base de su utilidad para evaluar la enorme diversidad de la evidencia que se puede hallar en relación a una determinada hipótesis científica (Stegenga, 2014). Los ensayos controlados aleatorizados (RCT) suelen ocupar las posiciones más destacadas de las jerarquías evidenciales, mientras que otros métodos como los estudios de cohorte, los de laboratorio o los estudios de casos ocupan lugares inferiores. La jerarquía evidencial conduce al monismo metodológico cuando se defiende la preeminencia del RCT por su capacidad para producir evidencia de gran precisión y validez interna. Desde este enfoque, el RCT se concibe como una metodología superior al resto porque, mediante la aleatorización y el ciego, puede minimizar los

* roberto_lopez_30@hotmail.com

sesgos que alterarían los resultados del estudio (Jukola, 2019; Lichtenstein et al., 2021). En consecuencia, los otros diseños de estudios se considerarían válidos y relevantes en la medida en que emulen los RCT (Vandenbroucke, Broadbent, & Pearce, 2016).

El monismo metodológico ha sido criticado por numerosos autores (Jukola, 2019; Reiss, 2015; Richardson, 2012; Stegenga, 2014; Vandenbroucke, Broadbent, & Pearce, 2016). Si bien los RCT ideales podrían ser aceptables para generar evidencia científica relevante, en la práctica pueden presentar numerosas dificultades y limitaciones. En primer lugar, se ha señalado que la aleatorización no garantiza la minimización de sesgos, ya que los factores de confusión podrían estar distribuidos de manera no equitativa entre el grupo de intervención y el de control. Este proceso solo podría eliminar la posibilidad de sesgos si se repitiera infinitamente. Por lo tanto, la aleatorización no convertiría a los RCT en un método epistémicamente superior al resto. Después, se ha destacado el problema de la validez externa: la extrapolación de los resultados de los RCT a individuos que se encuentren en contextos diferentes a los del estudio puede ser difícil o imposible. Aunque un ensayo pruebe una determinada eficacia en un entorno experimental, su potencial de generalización podría ser limitado.

Frente al enfoque monista, los críticos proponen un pluralismo metodológico que toma en consideración la totalidad de la evidencia científica –independientemente del método del que proceda– y el contexto en el que se esta se halla (Reiss, 2015; Richardson, 2012; Stegenga, 2014; Vandenbroucke, Broadbent, & Pearce, 2016). Rechazan que las metodologías científicas generen evidencia con una calidad inherente. Los pluralistas arguyen que en ámbitos como la medicina, la epidemiología o la nutrición, la práctica investigadora no se reduce a los RCT, sino que se utilizan diferentes tipos de evidencia en la evaluación de hipótesis. En este sentido, las relaciones causales se podrían llegar a establecer mediante estudios con diseños diferentes. Aunque un estudio *in vitro* o un estudio de cohorte no puedan probar causalidad individualmente, sería posible llegar a tal conclusión con la valoración conjunta de la evidencia.

Discusión y conclusión

Se analiza la controversia metodológica en la ciencia reguladora a partir del estudio de dos casos: la regulación de las declaraciones de salud y la regulación de las vacunas. Se ha defendido que el nivel de certidumbre en los datos requeridos para la toma de decisiones reguladoras debe tener en cuenta qué se está regulando y en qué contexto (Biesalski et al., 2011; Blumberg et al., 2010; Jukola, 2019). El nivel de prueba exigido debe depender de los riesgos que se asumen al aprobar o rechazar cada producto o estrategia, por lo que debe tomar en consideración sus consecuencias últimas. En este sentido, la regulación de las vacunas contra la COVID-19 expone una estrategia donde determinada evidencia se considera no concluyente para que una vacuna sea aprobada, pero suficiente para que sea

autorizada para su uso con el fin de alcanzar objetivos de salud pública. La evidencia no concluyente –con un nivel de fiabilidad menor a la producida por los RCT correctamente diseñados, desarrollados y finalizados– puede fundamentar la toma de decisiones reguladoras sobre la base de valores no epistémicos.

Referencias bibliográficas

- [1] Biesalski, H. K., P. J. Aggett, R. Anton, P. S. Bernstein, J. Blumberg, R. P. Heaney, ... I. Zöllner (2011). “26th Hohenheim Consensus Conference, September 11, 2010 scientific substantiation of health claims: evidence-based nutrition”, *Nutrition*, 27(10 Suppl.), pp. S1-S20. doi:10.1016/j.nut.2011.04.002
- [2] Blumberg, J., R. P. Heaney, M. Huncharek, T. Scholl, M. Stampfer, R. Vieth, ... S. H. Zeisel (2010). “Evidence-based criteria in the nutritional context”, *Nutrition Reviews*, 68(8), pp. 478-484. doi:10.1111/j.1753-4887.2010.00307.x
- [3] Jukola, S. (2019). “On the evidentiary standards for nutrition advice”, *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 73, pp. 1-9. doi:10.1016/j.shpsc.2018.05.007
- [4] Lichtenstein, A. H., K. Petersen, K. Barger, K. E. Hansen, C. A. M. Anderson, D. J. Baer, ... N. R. Matthan (2021). “Perspective: design and conduct of human nutrition randomized controlled trials”, *Advances in Nutrition*, 12(1), pp. 4-20. doi:10.1093/advances/nmaa109
- [5] Richardson, D. P. (2012). “Preparing dossiers: strength of the evidence and problems of proof”, *Proceedings of the Nutrition Society*, 71(1), pp. 127-140. doi:10.1017/S002966511100317X
- [6] Reiss, J. (2015). “A pragmatist theory of evidence”, *Philosophy of Science*, 82(3), pp. 341-362. doi:10.1086/681643
- [7] Stegenga, J. (2014). “Down with the hierarchies”, *Topoi*, 33, pp. 313-322. doi:10.1007/s11245-013-9189-4
- [8] Vandembroucke, J. P., A. Broadbent y N. Pearce (2016). “Causality and causal inference in epidemiology: the need for a pluralistic approach”, *International Journal of Epidemiology*, 45(6), pp. 1776-1786. doi:10.1093/ije/dyv341

Políticas epistémicas y controversias en ciencia reguladora

José Luis Luján*
University of the Balearic Islands

Oliver Todt**
University of the Balearic Islands

Short Abstract

En este trabajo se presenta un modo de analizar la relación entre controversias científicas y controversias públicas relacionadas con la regulación de aplicaciones tecnológicas. Nuestra propuesta consiste en identificar el vínculo entre objetivos reguladores y políticas epistémicas. Concluimos que la adecuación de las políticas epistémicas depende de los objetivos de la regulación. No es posible, por tanto, formular políticas epistémicas neutrales respecto de los objetivos reguladores.

Introducción

Desde finales de los años 60 del siglo XX las aplicaciones tecnológicas y/o sus efectos son objeto de la regulación por parte de gobiernos y parlamentos. Ejemplos de este tipo de regulaciones son las relativas a los riesgos de los productos tóxicos o las regulaciones medioambientales. En general, se trata de normativas cuyo objetivo suele ser la protección de la salud pública y/o del ambiente.

Tanto el establecimiento de estas regulaciones como su posterior evaluación requieren del uso de conocimiento científico. Es en este contexto en el que se consolida la denominada ciencia reguladora. Se trata de una actividad cuyo objetivo es en general asesorar la toma de decisiones en relación con la elaboración y evaluación de regulaciones relacionadas con aplicaciones tecnológicas.

La ciencia reguladora está relacionada con la regulación de numerosas aplicaciones tecnológicas: la biotecnología, los alimentos funcionales, los productos químicos, los fármacos, etc. Se trata por tanto de una actividad científica directamente relacionada con importantes debates sociales, intereses económicos y posiciones políticas. Por lo tanto, una

* j.l.lujan@uib.es

** oliver.todt@uib.es

actividad científica envuelta generalmente en controversias sociales que acaban teniendo también una contrapartida científica.

En las controversias en ciencia reguladora se plantean principalmente cuestiones metodológicas y metametodológicas. Las cuestiones metodológicas son relativas a las reglas de inferencia, los modelos de extrapolación, etc. Las cuestiones metametodológicas, por su parte, son relativas al estándar de prueba, las jerarquías evidenciarias, etc. Proponemos el uso del término ‘políticas epistémicas’ para referirnos a distintas combinaciones de opciones metodológicas y metametodológicas. Las políticas epistémicas determinan el tipo de conocimiento pertinente estableciendo la carga de la prueba, fijan los requisitos de justificación de las afirmaciones de conocimiento mediante el estándar de prueba, y optan por una jerarquía evidencial particular, especificando el valor de distintos tipos de información científica.

Políticas epistémicas y controversias científicas

Las controversias en ciencia reguladora pueden conceptualizarse como controversias entre políticas epistémicas alternativas. Esto es, como propuestas alternativas de estándares de prueba, jerarquías evidenciarias, definiciones de las relaciones causales, etc. Estas políticas epistémicas están a su vez relacionadas con distintos objetivos reguladores, lo que establece una conexión entre las controversias científicas y las controversias públicas (sociales y políticas) en las que habitualmente se ve envuelta la ciencia reguladora.

Nuestros casos de estudio son relativos a la evaluación de riesgos (productos químicos potencialmente tóxicos) y la evaluación de beneficios (evaluación de las propiedades saludables de los alimentos). Las controversias científicas en ambos casos están relacionadas con dos políticas epistémicas alternativas. La principal diferencia entre estas políticas epistémicas es el estándar de prueba. Tanto en la evaluación de riesgos como en la evaluación de beneficios encontramos la defensa de estándares de prueba más permisivos y otros más estrictos (Cranor 1995; Hendrickx 2013) Los estándares de prueba son el componente esencial de las políticas epistémicas. Esto es así porque influyen en muchos otros aspectos de la investigación.

De estándares de prueba (o evidenciarios) se puede hablar en un sentido restrictivo o en un sentido amplio. En un sentido restrictivo, la expresión estándares de prueba hace referencia al nivel y el tipo de evidencia que se requiere para justificar una afirmación empírica. En un sentido más amplio, los estándares de prueba no sólo se relacionan con la elección entre hipótesis alternativas, sino también con el proceso completo de generación de conocimiento sobre riesgos y/o beneficios. En el desarrollo de la investigación, decisiones tales como fijar el nivel de significatividad estadística, establecer criterios para aceptar o rechazar datos (caracterización de los datos), seleccionar los métodos para interpretar la evidencia, son decisiones vinculadas a los estándares de prueba (Douglas 2009).

Políticas epistémicas y controversias públicas

Las controversias públicas en torno a la regulación de la tecnología están relacionadas con los objetivos de la regulación. En el caso de la evaluación de riesgos estos objetivos son la protección de la salud, la protección del entorno, la promoción de la innovación, etc. En relación con los beneficios de los alimentos encontramos objetivos como la protección del consumidor frente a información incierta, la promoción de la salud mediante la dieta, etc.

El estudio de la relación entre la controversia científica y la controversia pública es posible analizando la conexión entre objetivos reguladores y políticas epistémicas (Andreoletti y Teira 2019; Hansson 2020; Osimani 2014). Así, en la evaluación de riesgos la protección del entorno y de la salud se relaciona con una política epistémica más permisiva (en el sentido de más cautelosa o precaucionaria), mientras que la política epistémica con un estándar de prueba más estricto se relaciona con priorizar la innovación. En el caso de la evaluación de beneficios un estándar más estricto está vinculado con la protección de los consumidores frente a posible información incierta. La promoción de la innovación en alimentación está relacionada con una política epistémica con un estándar más tolerante.

Conclusión

El estudio de las controversias reguladoras muestra una relación entre las políticas epistémicas y los objetivos reguladores. Esta relación supone un límite a la posibilidad de identificar una política epistémica que sirva para cualquier conjunto de objetivos reguladores. La adecuación de una política epistémica depende entonces de los objetivos reguladores: la elección de los estándares de prueba, la jerarquía evidenciaria, reglas de inferencia, modelos de extrapolación, etc., está relacionada con los objetivos reguladores.

Referencias

- [1] Andreoletti, M. y Teira, D. (2019). “Rules versus Standards: What Are the Costs of Epistemic Norms in Drug Regulation?” *Science, Technology and Human Values*, 44 (6): 1093-1115.
- [2] Cranor, C. (1995). “The Social Benefits of Expedited Risk Assessment”. *Risk Analysis*, 15 (4): 353-358.
- [3] Douglas, H. (2009). *Science, Policy, and the Value-Free Ideal*. Pittsburgh: University of Pittsburgh Press.

- [4] Hansson, S.O. (2020). “Values in pharmacology”. In *Uncertainty in pharmacology. Epistemology, methods, and decisions*, eds. LaCaze, A. and Osimani, B., 375-396. Cham: Springer.
- [5] Hendrickx, K. (2013). “Rivaling evidence-bases and politics in regulatory science”. *Food, Science & Law*, vol. 4, [Http://hdl.handle.net/2268/162196](http://hdl.handle.net/2268/162196)
- [6] Osimani, B. (2014). “Safety vs. efficacy assessment of pharmaceuticals: Epistemological rationales and methods”, *Preventive Medicine Reports* 1: 9-13.

A Pricean Foundation for Cultural Evolutionary Theory

Victor J. Luque*
University of Valencia

Lorenzo Baravalle**
Universidade de Lisboa

Formulated at the beginning of the 1970s by population geneticist George Price (1970, 1972), the *Price equation* is, nowadays, considered by some authors the key for a generalised theory of evolution (Rice 2004; Frank 2012 and 2017; Luque 2017). In this paper, we shall argue that the Price equation – or, at least, a version of it – is the *fundamental law* of cultural evolutionary theory. In order to support this claim, we shall consider two issues. The first one is genuinely theoretical. Cultural evolutionists have identified, throughout the last decades, a set of evolutionary factors responsible for cultural change. We tentatively put forward the following taxonomy (inspired in Mesoudi 2011 and Baravalle 2019) of the possible causes of cultural evolution:

- *Selective forces*: natural selection, cultural selection due to content and context-based biases.
- *Random forces*: genetic and cultural drift.
- *Transmission biases*: guided variation, random mutation, blending transmission.
- *Migration*: demic migration and cultural migration.

The question is: may the Price equation properly represent the causes of cultural evolution? The original Price equation does not say anything *specific* about cultural evolutionary processes. In order to support the claim that the Price equation can properly represent all the causes of cultural evolution, we consider two attempts to reformulate the equation so as to make it more explicitly suitable to account for cultural evolutionary dynamics. The first reformulation is by El Mouden et al. (2014):

$$\Delta\bar{z} = Cov(c, z) + E_c(\Delta z)$$

where $\Delta\bar{z}$ is the change in the average value of a cultural trait z . The first term on the right-hand side represents how much a cultural trait z and cultural fitness, c , covary; and

* vicluqmar@gmail.com

** lorenzo_baravalle@yahoo.it

$E_c(\Delta z)$ represents the transmission bias in cultural change, weighted by cultural fitness c . Drawing on this work, Aguilar and Akçay (2018) have presented a similar formalism to model cultural change. The difference with the previous cultural Price equation is that Aguilar and Akçay include both genetic and cultural inheritance systems in their formalism. Thus, Aguilar and Akçay present the change of average value of a behavioural trait (represented as a phenotypic trait, p) as follows:

$$\Delta \bar{p} = \frac{1}{\bar{w}} \text{cov}(w_i, g_i) + \frac{1}{\bar{w}} \text{cov}(s_i, c_i) + \langle \Delta g_j \rangle + \langle \Delta c_j \rangle$$

where the first two terms in the right-hand side measure the covariance between reproductive fitness w_i and a genotype g_i , and cultural fitness s_i and a cultural type c_i (respectively). The last two terms measure the transmission bias in genetic, Δg_j , and cultural, Δc_j , inheritance. Both equations are able to represent the several cultural forces delineated above.

The second issue that we shall consider says about the *empirical* import of the Price equation. Even accepting that the formal descriptions of cultural evolutionary forces can be accommodated within the Price equation, this does not entail that we can satisfactorily employ the equation for explanatory or predictive goals. We argue that if, in a specific scenario, we are able to identify at least some evolutionary forces, then we can employ the Price equation to provide hypotheses about the presence of other causes of evolution and its characteristics. More specifically, we argue that:

- a. the Price equation is explanatory insofar as, when considered in relation with specific cultural evolutionary scenarios, it provides heuristic guide about the possible causes of the phenomena under study; and
- b. it is precisely because this that the Price equation (or, better, a specific cultural version of the Price equation, like Aguilar and Akçay's equation) is a good candidate to be the fundamental law, or the *guiding-principle* (Moulines 1984; Díez and Lorenzano 2013), of cultural evolutionary theory.

We illustrate this point through three main case studies (MacCallum et al. 2012; Aguilar and Akçay 2018; and Gong et al. 2012).

Finally, we fully characterise the cultural version of the Price equation as the fundamental law of cultural evolutionary theory, by employing some conceptual resources of

metatheoretical structuralism (Balzer et al. 1987). According to metatheoretical structuralism, mature scientific theories are typically composed of a set of theoretical elements arranged in an inverted-tree hierarchy, also called *theory net*. The theoretical elements are set-theoretic structures identifying the models of the theory. At the top of the hierarchy we usually find a single theoretical element, including some very general assertion about the way in which the theory account for a class of phenomena: this is the fundamental law. In the case of cultural evolutionary theory, the cultural Price equation is the fundamental law or guiding principle since establishes “what the theory is about” by providing a very general and comprehensive schema for explaining a class of phenomena. Thus, the cultural Price equation provides a very abstract and general statement about the kind of factors that can produce changes in the system under study (in this case, frequencies of cultural variants through time). There is no cultural evolutionary change without these forces acting on a population of traits and, at the same time, there are no other forces acting in the evolutionary process besides them.

The theoretical elements that branch beneath the top theoretical element are the *specialisations* of the top element: they specify regularities concerning some subclasses of the phenomena accounted for the theory. While the Price equation describes the net change in the diffusion of a trait as the product, or effect, of forces, the specialisations of the theory provide overall causal characterisations of each force, and these characterisations allow, in their turn, empirical descriptions of the processes acting in cultural change. A precise characterisation of the overall causal features of the evolutionary factors is what enables reliable hypotheses about concrete scenarios (as in the three main case studies mentioned above).

In summary, we argue that cultural evolutionary theory is an evolutionary theory because the forces of cultural evolution (as usually conceptualised in dual-inheritance theory) can be easily plugged in a theoretical structure in which a specific cultural version of the Price equation is the fundamental law. As well as in other domains of evolutionary theory, the Price equation works in cultural evolutionary theory as a heuristic principle, providing a guide to the formulation of explanations of specific phenomena, and a unifying framework to think cultural evolutionary processes collectively. In doing so, we also tentatively depict the overall structure of cultural evolutionary theory. We think that this is an important task, since – if attained – it would be evidence that cultural evolutionary theory is somehow unitarily grounded; a claim that has recently been challenged (e.g., Lewens 2015).

References

- [1] Aguilar, E. and Akçay, E. (2018). “Gene-Culture Co-Inheritance of a Behavioral Trait”. *The American Naturalist*, 23, pp. 311-320.
- [2] Balzer, W., Moulines, C. U. and Sneed, J. D. (1987). *An Architectonic for Science. The Structuralist Program*. Dordrecht: Reidel.
- [3] Baravalle, L. (2019). “Cultural Evolutionary Theory as a Theory of Forces”. *Synthese*, <https://doi.org/10.1007/s11229-019-02247-0>.
- [4] Díez, J. and Lorenzano, P. (2013). “Who Got What Wrong? Fodor and Piattelli on Darwin: Guiding Principles and Explanatory Models in Natural Selection”. *Erkenntnis*, 78, pp. 1143-1175.
- [5] El Mouden, C., André, J.-B., Morin, O. and Nettle, D. (2014). “Cultural Transmission and the Evolution of Human Behaviour: A General Approach Based on the Price Equation”. *Journal of Evolutionary Biology*, 27, pp. 231-241.
- [6] Frank, S. A. (2012). “Natural selection. IV. The Price equation”. *Journal of Evolutionary Biology*, 25, pp. 1002-1019.
- [7] Frank, S. A. (2017). “Universal Expressions of Population Change by the Price Equation: Natural Selection, Information, and Maximum Entropy Production”. *Ecology and Evolution*, 7, pp. 3381-3396.
- [8] Gong, T., Shuai, L., Tamariz, M. and Jäger, G. (2012). “Studying Language Change Using Price Equation and Pólya-urn Dynamics”. *PLoS ONE*, 7, e33171.
- [9] Lewens, T. (2015). *Cultural Evolution*. Oxford: Oxford University Press.
- [10] Luque, V. J. (2017). “One Equation to Rule Them All: A Philosophical Analysis of the Price Equation”. *Biology and Philosophy*, 32, pp. 97-125.
- [11] MacCallum, R. M., Mauch M., Burt A. and Leroi A. M. (2012). “Evolution of Music by Public Choice”. *PNAS*, 109, pp. 12081-6.
- [12] Mesoudi, A. (2011). *Cultural Evolution: How Darwinian Theory can Explain Human Culture and Synthesize the Social Sciences*. Chicago: The University of Chicago Press.
- [13] Moulines, C. U. (1984). “Existential Quantifiers and Guiding Principles in Physical Theories”. In J. J. E. Gracia, E. Rabossi, E. Villanueva and M. Dascal (eds). *Philosophical Analysis in Latin America*. Dordrecht: Reidel, pp. 173-198.
- [14] Price, G. R. (1970). “Selection and Covariance”. *Nature*, 227, pp. 520-521.

- [15] Price, G. R. (1972). "Extension of Covariance Selection Mathematics". *Annals of Human Genetics*, 35, pp. 485-490.
- [16] Rice, S. H. (2004). *Evolutionary Theory: Mathematical and Conceptual Foundations*. Sunderland, MA: Sinauer Associates.

¿Es posible ser pragmatista en ciencia?

Una caracterización deflacionaria de la comprensión científica

Mariano Martín Villuendas*
Universidad de Salamanca

En los últimos años, la noción de comprensión se ha convertido en uno de los principales focos de discusión dentro de los estudios dedicados a la epistemología y filosofía de la ciencia. Dentro de este último ámbito, han surgido innumerables propuestas acerca de cómo debería ser entendida, cuáles deberían ser sus condiciones de surgimiento o cuál debería ser su estatus dentro de la empresa científica. A pesar de haber sido un concepto ampliamente debatido, no existe en la actualidad un acuerdo unánime en torno a estas últimas cuestiones. Tanto es así que la discusión se ha polarizado en dos posiciones marcadamente diferenciadas.

Por un lado, se encontrarían los autores que han abordado el problema de la comprensión valiéndose de las herramientas conceptuales propias de la epistemología tradicional, esto es, empleando los conceptos epistemológicos de verdad y conocimiento (Frigg y Nguyen, 2019; Greco, 2014; Grimm, 2012; Kahlifa, 2017; Kelp, 2015; Kuorikoski y Ylikoski, 2015; Kvanvig, 2009; Rice, 2017; Strevens, 2008; Trout, 2007). Los análisis llevados a cabo por estos autores parten de tres presuposiciones básicas. La primera es que la verdad constituye el objetivo epistémico último de la empresa científica. La segunda es que el conocimiento, entendido como creencia verdadera y justificada, se articula a través de la explicación. La tercera es que existe un vínculo estrecho entre la explicación y la comprensión, en tanto que la segunda depende de la primera. Dado que el conocimiento verdadero se estructura por medio de explicaciones y puesto que son estas las que originan la comprensión, entonces este último concepto debe analizarse en términos de verdad.

Esta primera posición ha tenido, sin embargo, serios problemas a la hora de acomodar una característica fundamental de la empresa científica: el extendido y variado empleo de idealizaciones y abstracciones. Diversos estudios (Jones y Cartwright, 2005; Morgan y Morrison, 1999; Morrison, 2015; Potochnik, 2017; Rice, 2015, 2017; Suárez, 2009) han demostrado que las idealizaciones, además de afectar a los supuestos centrales de los modelos, tienen un carácter marcadamente holístico –distorsionan las interacciones que

* mmvilluendas@gmail.com

se establecen entre los elementos del sistema. Un claro ejemplo de ello lo constituyen los modelos de optimización en biología o los modelos climáticos. Se ha demostrado que son precisamente esta clase de distorsiones, dependientes de los intereses pragmáticos de investigación, las que permiten a los modeladores emplear técnicas de modelado matemático con las que poder acceder a explicaciones que, de otro modo, serían inaccesibles (Parker, 2020a). De igual manera, esta primera postura, al hacer depender de manera directa la comprensión de la explicación, ha obviado la enorme pluralidad y heterogeneidad de intereses cognitivos que pueblan la práctica científica, siendo incapaz, por tanto, de dar cuenta de una amplia clase de modelos científicos –climáticos, taxonómicos o computacionales– (Durán y Formanek, 2018; Humphreys, 2004; Lenhard, 2019).

A fin de afrontar estas dificultades, diversas autoras y autores (de Regt, 2017; Elgin, 2017; Leonelli, 2013; Potochnik, 2017) han apostado por realizar una caracterización alternativa de la comprensión científica. La comprensión, argumentan, debería ser analizada tomando como punto de referencia los intereses pragmáticos de los agentes cognitivos, sus capacidades manipulativas y las particularidades del contexto de investigación considerado. Esto es, la comprensión debería ser analizada con independencia de cualquier consideración relativa a la verdad. Analizar la validez de las diversas propuestas pragmáticas se ha convertido en uno de los principales objetivos dentro de la literatura concerniente a la comprensión. El motivo es claro: si el valor de la comprensión supera al del conocimiento verdadero y si el objetivo de la ciencia es la comprensión y no la verdad. ¿No supone esto poner en cuestión el edificio epistemológico tradicional sobre el cual se fundamenta la práctica científica? (Frigg y Nguyen, 2019).

El objetivo de la presente comunicación será argumentar a favor de esta última posición. Para ello, estructuraré la misma en dos secciones.

En la primera, analizaré las virtudes y defectos de una de las principales propuestas pragmáticas de la actualidad, la teoría delineada por Henk de Regt (2017). Mostraré cómo la postura de de Regt afronta dos problemas esenciales que es preciso solventar si se pretende articular una noción de comprensión pragmática viable. En primer lugar, mostraré en qué medida esta teoría es incapaz de acomodar modelos cuyo funcionamiento es parcialmente independiente de la teoría (Lenhard, 2019). Ilustraré este primer punto a través del modelado climático (Frigg *et al.*, 2015a, 2015b; Parker, 2020b). En segundo lugar, mostraré cómo, al ligar el concepto de comprensión al de explicación, la teoría de de Regt es incapaz de dar cuenta de la amplia variedad de objetivos cognitivos que pueblan la práctica científica – predicción, clasificación, enseñanza, unificación, entre otros.

En la segunda, con el objetivo de afrontar las dificultades que afectan a la postura de de Regt y de mostrar la viabilidad de la propuesta pragmática, ofreceré una noción alternativa e innovadora de comprensión. La noción deflacionaria y minimalista de

comprensión propuesta se basará en cinco condiciones necesarias: Se dice que alguien tiene comprensión cuando (1) existe la condición de posibilidad para poder (2) ejercer una(s) determinada(s) capacidad(es) (3) fundada(s) sobre los estándares de aceptación de una determinada comunidad de agentes, con el objetivo explícito de satisfacer, de manera adecuada, (4) determinados intereses/necesidades cognitivas (5) de manera eficiente y coherente (lógica, empírica, metodológica y epistémicamente). Mostraré cómo esta noción de comprensión es capaz de afrontar los problemas que aquejan tanto a la postura veritista –uso de idealizaciones– como a la postura pragmática de de Regt –uso de modelos parcialmente independientes de la teoría y variedad de objetivos cognitivos.

Referencias

- [1] De Regt, Henk (2017). *Understanding Scientific Understanding*. New York: Oxford University Press.
- [2] Durán, Juan y Formanek, Nico (2018). Grounds for Trust: Essential Epistemic Opacity and Computational Reliabilism. *Minds and Machines*, 28(4), 645-666.
- [3] Elgin, Catherine (2017). *True Enough*. Cambridge: The MIT Press.
- [4] Frigg, Roman, Thompson, Erica y Wendl, Charlotte (2015). Philosophy of Climate Science Part I: Observing Climate Change. *Philosophy Compass*, 10(12), 953-964.
- [5] Frigg, Roman, Thompson, Erica y Wendl, Charlotte (2015). Philosophy of Climate Science Part II: Modelling Climate Change. *Philosophy Compass*, 10(12), 965-977.
- [6] Frigg, Roman y Nguyen, James (2019). Mirrors without warnings. *Synthese*.
- [7] Greco, John (2014). Episteme: Knowledge and Understanding. En Kevin Timpe y Craig Boyd (Eds.), *Virtues and Their Vices* (285-303). Oxford: Oxford University Press.
- [8] Grimm, Stephen (2012). The value of understanding. *Philosophy Compass*, 7(2), 103–117.
- [9] Humphreys, Paul (2004). *Extending Ourselves. Computational Science, Empiricism, and Scientific Method*. New York: Oxford University Press.
- [10] Jones, Martin y Cartwright, Nancy (eds.) (2005). *Idealization XII: Correcting the Model. Idealization and Abstraction in the Sciences*. Amsterdam: Rodopi.
- [11] Kelp, Christoph (2015). Understanding Phenomena. *Synthese*, 192, 3799-3816.
- [12] Khalifa, Kareem (2017). *Understanding, Explanation, and Scientific Knowledge*. Cambridge: Cambridge University Press.

- [13] Kuorikoski, Jaakko y Ylikoski, Petri (2015). External representations and scientific understanding. *Synthese* 192, 3817-3837.
- [14] Kvanvig Jonathan (2009). Responses to Critics. En Adrian Hadock, Alan Millar y Duncan Pritchard (Eds.), *Epistemic Value* (pp. 339-353). Oxford: Oxford University Press.
- [15] Lenhard, Johannes (2019). *Calculated Surprises. A Philosophy of Computer Simulation*. New York: Oxford University Press.
- [16] Leonelli, Sabina (2013). Understanding in biology: The impure nature of Biological Knowledge. En Henk de Regt, Sabina Leonelli y Kai Eigner (Eds.), *Scientific Understanding. Philosophical Perspectives* (pp. 189-210). Pittsburgh: University of Pittsburgh Press.
- [17] Morgan, Mary y Morrison, Margaret (eds.) (1999). *Models as Mediators. Perspectives on Natural and Social Science*. Cambridge: Cambridge University Press.
- [18] Morrison, Margaret (2015). *Reconstructing Reality: Models, Mathematics, and Simulations*. Oxford: Oxford University Press.
- [19] Parker, Wendy (2020a). Model Evaluation: An Adequacy-for-Purpose View. *Philosophy of Science*, 87, 457-477.
- [20] Parker, Wendy (2020b). Evidence and Knowledge from Computer Simulation. *Erkenntnis*.
- [21] Potochnik, Angela (2017). *Idealization and the Aims of Science*. Chicago: The University of Chicago Press.
- [22] Rice, Collin (2015). Factive Scientific Understanding without Accurate Representation. *Biology & Philosophy*, 31, 81-102.
- [23] Rice, Collin (2017). Idealized models, holistic distortions, and universality. *Synthese*, 195, 2795-2819.
- [24] Strevens, Michael (2008). *Depth: An Account of Scientific Explanation*. Cambridge: Harvard University Press.
- [25] Trout, J. D. (2007). The Psychology of Scientific Explanation. *Philosophy Compass*, 2(3), 564-591.

Programas de Lakatos e investigación antiviral contra el SARS-CoV-2

David Álvaro Martínez*
Universidad Complutense de Madrid

Resumen

La pandemia de la COVID-19 causada por un nuevo coronavirus, el SARS-CoV-2, ha vuelto imprescindible el desarrollo de nuevas vacunas y métodos de investigación para su control y erradicación. Esto ha supuesto la necesidad de adaptar y poner en marcha programas antivirales a la altura de las circunstancias. En este trabajo trataremos de esbozar las bases que caracterizan estos programas, valiéndonos de la metodología formulada por Imre Lakatos en *La metodología de los programas de investigación científica* (1978). En un contexto donde el escepticismo ante las vacunas puede hacer fracasar la implementación de dichos programas, es importante entender y transmitir a la población si los efectos secundarios de las vacunas, por ejemplo, afectan al núcleo y la efectividad de los mismos. Para ello, tomaremos como referencia las estrategias heurísticas propuestas por la Agencia Europea de Medicamentos (EMA), organismo encargado de la evaluación y autorización de medicamentos en la Unión Europea.

Desarrollo

En los años setenta, Thomas Kuhn discutió ampliamente las nociones progresistas de la ciencia formuladas por su coetáneo Karl Popper. Argumentaba Kuhn que los paradigmas científicos no se encuentran más cerca de la verdad unos de otros, sino que pertenecen a períodos diferentes de la ciencia e incommensurables entre sí. La metodología formulada por Lakatos que vamos a utilizar en este trabajo tratará de conciliar, por un lado, los postulados progresistas de Popper y, por otro, las nociones relativistas de Kuhn, en lo que se conoce como PIC (Programas de Investigación Científica). En un contexto actual donde el escepticismo ante las vacunas puede hacer fracasar la implementación de dichos programas, es importante entender y transmitir a la población si los efectos secundarios de las vacunas, por ejemplo, afectan al núcleo y la efectividad de los mismos.

Según Lakatos, un PIC supone una unidad descriptiva que facilita el abordaje teórico, la sistematización y la socialización de logros científicos. Este se compone de elementos

* davida11@ucm.es

relacionados según un plan inicial común o “núcleo firme”, es decir, principios y enunciados axiomáticos que sirven como base para realizar hipótesis posteriores o experimentos auxiliares. Dichas hipótesis forman un “cinturón protector” o revestimiento del núcleo, que recibirá el “impacto” de las contrastaciones y será sometido a pruebas de testeo y falsación, ajustándose y reajustándose para defenderlo (*cf.* Lakatos, 1989, 66). No obstante, el núcleo de un programa no tiene por qué ser irrefutable a toda costa, y puede “derrumbarse” o abandonarse cuando deje de anticipar hechos nuevos. De ser así, estaríamos ante un programa fallido que conducirá, en lugar de a cambios progresivos en las redes de problemas, a cambios regresivos. Además, todo PIC para Lakatos conlleva una determinada heurística; es decir, conjuntos de estrategias que ofrezcan rutas de investigación interdependientes a seguir (heurística positiva), frente a otras que deben evitarse (heurística negativa).

Teniendo en cuenta las críticas favorables y opuestas planteadas a la propuesta de Lakatos (*cf.* Cova, 2005; Fabbricatore, 2011), trataremos de seguir su metodología no buscando sino entender y analizar, desde una perspectiva rigurosa más, cómo se han formulado las distintas hipótesis que han dado lugar a cambios progresivos en la elaboración de vacunas antivirales contra el SARS-CoV-2. Para ello, nos apoyaremos en los trabajos realizados sobre virología de Urququi, S. & Ossa, J. (eds.) (2008), Adamson (ed.) (2020) y Dulbecco & Ginsberg (1985). A partir de estos trabajos, podemos reconstruir el núcleo de un programa antiviral como compuesto por cuatro premisas básicas:

- 1) Un virus, siendo un agente infeccioso microscópico acelular –esto es, absolutamente dependiente de células para replicarse–, se clasifica según su ácido nucleico (DNA o RNA), su simetría (icosaédrica, helicoidal o indeterminada) y la presencia o ausencia de envoltura.
- 2) La replicación viral implica una exposición del material genético (adherencia, penetración o desnudamiento), una determinada expresión genética (síntesis de proteínas), la replicación del genoma, su morfogénesis y la liberación de la célula hospedera.
- 3) La respuesta inmunológica a la expansión viral puede ser innata (inespecífica) o adaptativa (específica); en función de ello se realizará la preparación de respuestas antivirales activas, las cuales deben inhibir el metabolismo patógeno y debilitar cepas posteriores.
- 4) La respuesta antiviral debe ser excretada sin metabolizarse en el organismo ni generar metabolitos tóxicos.

Siguiendo estas premisas podemos elaborar clasificaciones víricas básicas y de replicación, y preparar respuestas antivirales específicas. Es importante recalcar, como apunta Lakatos, que el núcleo del programa no nace ya dotado de toda su fuerza, sino que es producto de largos procesos preliminares de ensayo y error (*cf.* Lakatos, 1989, 67). Por

ejemplo, el tamaño de los virus, aunque hoy no se considera una de las características biológicamente más importantes, fue el parámetro utilizado durante mucho tiempo para su definición (*cf.* Urququi, S., Ossa, J., 2008).

Ahora bien, una vez constituido el núcleo y puesto en funcionamiento, este resulta “irrefutable” por propia decisión metodológica. Las nuevas anomalías, en este caso la irrupción pandémica por un nuevo coronavirus SARS-CoV-2 en humanos, provocará cambios en el cinturón protector del programa, aunque no en el núcleo, pues tanto el virus SARS-CoV-2 como sus respuestas antivirales siguen siendo identificables dentro del mismo.

Lo que nos servirá para contrarrestar las hipótesis falsadoras y teorías externas que emergen de los nuevos fenómenos será la elaboración de hipótesis y pruebas auxiliares, tal y como se aprecia en la *imagen 1*. En el caso del SARS-CoV-2, el cinturón protector deberá contrarrestar y explicar dificultades como el origen preciso de su transmisión (*imagen 2*), estableciendo hipótesis neutralizantes, siendo la más extendida la de transmisión zoonótica (Cui J. et al., 2019).

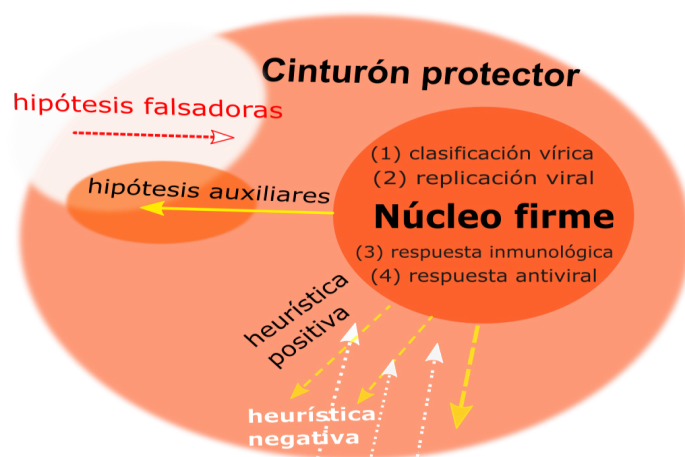


imagen 1

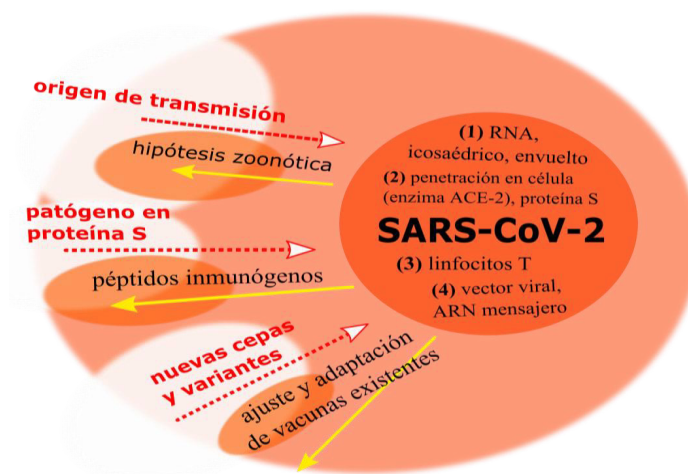


imagen 2

En la imagen 2 podemos apreciar ejemplos de hipótesis neutralizantes que, mediante la heurística positiva del programa, conforman el cinturón protector. Por ejemplo, para contrarrestar hipótesis potencialmente falsadoras como la codificación de los antígenos del patógeno (proteína S del virus), se llevarán a cabo experimentos y pruebas de testeo en bioinformática y proteómica para encontrar los péptidos inmunógenos (cf. Lozada-Requena, I., Núñez P. C., 2020). Decimos “potencialmente falsadoras” aquí porque, de no poder neutralizar la problemática que supone codificar los antígenos, el propio programa correría el riesgo de derrumbarse.

Por otro lado, ante fenómenos como las nuevas mutaciones o cepas, la hipótesis auxiliar más satisfactoria si atendemos a la literatura especializada (cf. Torres et al., 2021) será la de adaptar las vacunas ya existentes a cada variante. Todo ello no debilitaría, a priori, el núcleo del programa, pues sus premisas siguen funcionando y, de nuevo, es el cinturón protector el encargado de acomodar el impacto de las nuevas variables.

Si el programa consigue adaptarse, asimilando hechos futuros, hablaremos de un modelo antiviral exitoso o progresivo; si se torna regresivo, o incapaz de asimilar la evidencia empírica contraria, se convertirá en un modelo fallido. Ejemplos con respecto al SARS-CoV-2 de esto último serían, hasta el momento, programas como los de MSD y Sanofi & GSK, cuyas respuestas inmunitarias han resultado más deficientes que las de personas que contraen el virus de forma natural (cf. Griffin, 2021; Sanofi, 2020).

Para definir una metodología heurística fiable, tomaremos como referencia las estrategias marcadas por la Agencia Europea de Medicamentos (EMA). Una idea general y condensada de lo que supone este conjunto de reglas se refleja en la imagen 3.

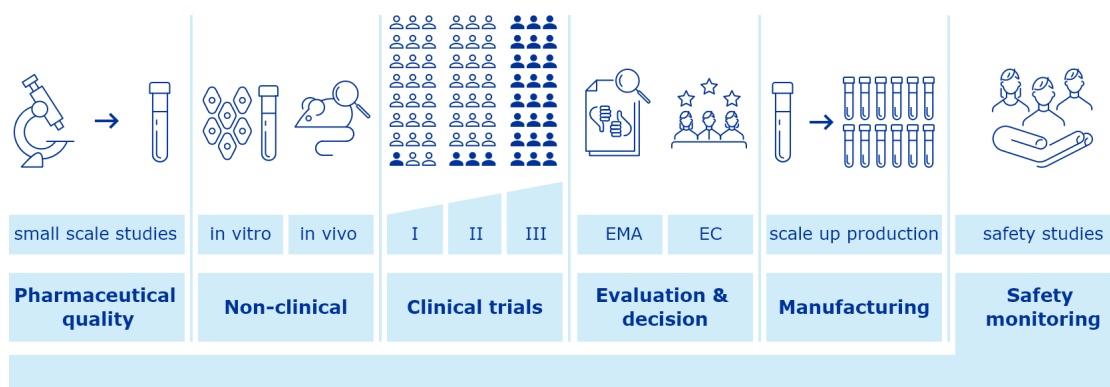


imagen 3. www.ema.europa.eu (cf. European Medicines Agency, 2018)

En los tres primeros pasos –(1) estudios a pequeña escala, (2) ensayos no clínicos *in vitro* e *in vivo*, (3) ensayos clínicos en individuos/grupos aislados y completos– las compañías y grupos de investigación operarían de manera independiente, buscando garantizar

la neutralidad y no otro interés que el desarrollo de la investigación. No obstante, la EMA, o cualquier organismo equivalente, puede proporcionar asesoramiento científico para potenciar un desarrollo más eficaz de los recursos científicos limitados, en beneficio de los pacientes.

Tras ello, solo un pequeño número de resultados será lo suficientemente prometedor como para ser probado en pacientes. Asimismo, solo una fracción de estos obtendrá resultados de estudio lo suficientemente buenos como para asegurar una calidad farmacéutica deseada y llegar al mercado. Más adelante, se evaluará y decidirá la viabilidad de los modelos (4) a través de la EMA y la EC (Comisión Europea). Una vez aprobados, se procederá a su manufacturación y producción a escala proporcional (5), seguida de un monitoreo de seguridad en los distintos pacientes (6).

Estas estrategias suponen un conjunto de reglas heurísticas que han sido probadas como altamente progresivas, y así lo corroboran los hechos (*cf.* AstraZeneca, Univ. Oxford 2020; Pfizer BioNTech 2020). No obstante, ello no significa que solo sean estas reglas las únicas posibles, pudiendo variar y reconfigurar el programa según las circunstancias demandadas. Todo ello sin perder de vista el enfoque eminentemente práctico del programa, de cara a resolver problemas y despejar interrogantes en tiempos de pandemia.

Bibliografía

- [1] Adamson, C. (ed.) (2020). *Antiviral Agents*, MDPI AG.
- [2] AstraZeneca, University of Oxford (2020). “Oxford University breakthrough on global COVID-19 vaccine” en www.ox.ac.uk
- [3] Cova, A. et al (2005). “Lakatos y los programas de investigación científica. Una opción para la organización investigativa nacional” en *Omnia*, 11
- [4] Cui J. et al. (2019). “Origin and evolution of pathogenic coronaviruses” en *Nat Rev Microbiol*
- [5] Dulbecco, R., Ginsberg H. (1985). *Virologia*, Piccin-Nuova Libreria.
- [6] European Medicines Agency (2018). “From laboratory to patient: the journey of a centrally authorized medicine” en *EMA*.
- [7] Fabbricatore, E. (2011). “Una valoración de la metodología de Imre Lakatos desde la práctica científica contemporánea” en *Revista Laguna*, 29.
- [8] Griffin, P. (2021). “Merck Shuts Down Covid Vaccine Program After Lackluster Data” en *Bloomberg*.

- [9] Lakatos, I. (1989). *La metodología de los programas de investigación científica*, Alianza Universidad (obra original publicada en 1978).
- [10] Lozada-Requena, I., Núñez P. C. (2020). “COVID-19: respuesta inmune y perspectivas terapéuticas” en *Rev Peru Med Exp Salud*, 7.
- [11] Pfizer BioNTech (2020). “Pfizer and BioNTech Achieve First Authorization in the World for a Vaccine to Combat COVID-19” en www.pfizer.com
- [12] Sanofi (2020), “Sanofi y GSK anuncian un retraso en su programa de vacuna COVID-19” en www.sanofi-pacifico-caribe.com
- [13] Torres et al. (2021). “Variantes y mutaciones del SARS-CoV-2” en *Proyecto PAIS*.
- [14] Urcuqui, S., Ossa, J. (eds.) (2008). *Principios de virología*, Universidad de Antioquia.

Health: a new conceptualization for the post-pandemic scenario. When philosophy of medicine meets environmental philosophy

Laura Menatti*

UMR 5319 Passages - École nationale supérieure d'architecture et de Paysage (ENSAP) Bordeaux

Cristian Saborido Alejandro**
UNED Madrid

During the last two years the COVID-19 pandemic has radically challenged the conceptualisation of health and the management of health care. To understand the importance of this major challenges on the definition of health, we propose a theoretical framework based on the collaboration between philosophy of science, philosophy of medicine and environmental philosophy. We aim to provide a better understanding of what health means in times of COVID-19 by analysing the environmental features implied.

The mainstream definition of health as proposed by the WHO – and used in medical education and medical epistemology – is: ‘a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity’ (WHO, 1948). We question this definition and we introduce a wider characterisation of health through different documents and supported by a theoretical analysis at the crossroad of philosophy of medicine and environmental philosophy. We will start from a 1984 WHO document in which health is defined as: “the extent to which an individual or group is able to realise aspirations and satisfy needs and to change or cope with the environment. Health is a resource for everyday life, not the objective of living; it is a positive concept, emphasising social and personal resources, as well as physical capacities” (WHO, 1984). This same definition was taken up in following documents, such as the Ottawa Charter (1986). Two important elements of this definition can be considered of relevance: cope and environment.

The conceptual relation between the environment and health has been present in the history of medicine since its origins. For instance, in the Hippocratic corpus (V BCE - II CE) the environment is correlated to several symptoms and conditions (Porter 1998; Bridgman and Gorsky 2012). Humoral pathology, based on the balance of the four humours (blood, phlegm, yellow bile and black bile), based on the Corpus Hippocraticum and

* laura.menatti@gmail.com

** cristian.saborido@fsof.uned.es

Galen's theory, remained a fundamental in medical epistemology and practice all over the Renaissance and until 1800 (Pormann 2019). Humour pathology was often accompanied by miasma theory, which was predominant until the introduction of germ theory and nowadays no longer accepted by the scientific community. Within the miasma theory, the environment was mostly reduced to air. Miasma theory was questioned by J. Snow during the epidemic of cholera in London at the end of 1800 and later substituted by germ theory by Pasteur and Koch (Bingham et al. 2004; Porter 1998: 428). The environment-health relationship has radically changed with the introduction of germ theory at the end of the 19th century. Since then, microbiological, ecological (considered as adaptive immune responses), immunological, and genetic theories have not only been complementary, but also overlapping (Casanova & Abel 2013). In the last years, the relationship between the environment and health has been discussed with respect to climate change. In 2009, the journal *Lancet*, together with University College London Institute for Global Health Commission (UCL) published a special issue on 'Managing the health effects of climate change' (Costello et. al 2009).

Furthermore, the environment appears in the long list of Social Determinants of Health (2008), together with the various conditions which have an impact on health, among them: governance, education, employment, social security etc. There is a long list of international and global official documents – such as the Millennium ecosystem assessment or the OneHealth approach – which calls for the importance of the environment in healthcare.

Yet a specific definition and analysis of health as related to 'environment' has not been provided in humanities or philosophy of medicine. In most theoretical proposals from philosophy of medicine, the environment is seen as the 'boundary conditions' to which the organism must adapt or it is considered as pathogenic, thus carrier of diseases. This has led philosophers of medicine to focus mainly on the possible health challenges posed by the environment. Moreover, the most widespread narratives about the recent pandemic have further emphasised the potentially pathogenic nature of the environment.

We thus propose to take into consideration the pathogenic effects of the environment, but also consider its salutogenic value and effects. In this paper we will show how both salutogenic and pathogenic dimensions are essential for a better understanding of the environment-organism relationship, with important theoretical and practical consequences for health and healthcare systems.

Salutogenesis is a medical framework developed by Antonovsky (1979) which questions the individualistic and disease-based approaches in medicine, and it is usually opposed to pathogenesis. It is focused on the origins of health and well-being, versus the aetiology of disease and the origins of risk factors. The salutogenic framework was used to describe the relationship between health, stress and coping. It has been developed to

foster health promotion approaches and programs able to move beyond risk factors and aimed at long-term health and well-being outcomes (Mittelmark et al. 2017). It focuses on resources for health and health-promoting processes, emphasising, for example, the role of prevention as well as the need of preventive medicine, or the importance of health education in medicine above risk factors.

With this in mind, we introduce the second important concept of the above mentioned definition of health of 1984, the notion of cope. We consider that the concept that better links the salutogenic and pathogenic effects of the environment on health comes from the notion of adaptivity, understood both from the perspective of an individual and as a social capacity. Adaptivity will be the conceptual tool to understand the pathogenic and salutogenic relationships between the environment and individuals.

Theoretical biology and philosophy of biology have deepened the analysis of the concept of adaptivity, in order to understand it as the relationship between the organism and its proximal environment (Di Paolo 2005; Moreno and Mossio 2015; Bich et al., 2016). More specifically, recent works in philosophy of biology have focused on investigating how, and by means of which mechanisms, organisms are able to compensate for variations (such for instance changes in the environment but also in their physiology). In this context, to analyse the complexity of the organism-environment relationship and to describe the capability to compensate for environmental variations, we think that the concept of regulatory adaptivity or adaptive regulation (Bich et al. 2016) could better serve our purpose. Adaptivity becomes thus not the mere passive viability of an organism or the capacity to remain stable to perturbations (that sometimes is reduced to homeostasis), but also the capability of the organism to modulate, and therefore change, its own dynamics by means of specialised mechanisms that are sensitive to perturbation.

This concept of individual adaptivity is suitable, according to us, to understand what adaptivity means to the environment and it is useful to explain both an individual and social scenario during the COVID-19 times.

References

- [1] Antonovsky, A. (1979). *Health, stress and coping*. San Francisco, CA: Jossey-Bass.
- [2] Berridge, V. & Gorsky, M. (Eds). (2012). *Environment, Health and History*. New York: Palgrave Macmillan.
- [3] Bich, L., Mossio, M., Ruiz-Mirazo, K. and Moreno, A. (2016). “Biological regulation: controlling the system from within”. *Biology & Philosophy*, 31(2), 237-265
- [4] Bingham, P., Verlander, N.Q., Cheal, M.J. (2004). “John Snow, William Farr and the 1849 outbreak of cholera that affected London: a reworking of the data highlights the importance of the water supply”. *Public Health*. 118(6):387-94.

- [5] Casanova, J. L., & Abel, L. (2013). "The genetic theory of infectious diseases: a brief history and selected illustrations". *Annual review of genomics and human genetics*, 14, 215-243.
- [6] Costello, A. et al. (2009). "Managing the health effects of climate change". *The Lancet*, Volume 373, Issue 9676, 1693-1733.
- [7] Di Paolo, E. (2005). "Autopoiesis, adaptivity, teleology, agency". *Phenom Cog Sci.* 4(4): 429-452.
- [8] Mittelmark, M.B, Sagy, S., Eriksson, M., Bauer, G.F., Pelikan, J.M., Lindström, B., Espnes, G.A. (2017) (Eds). *The Handbook of Salutogenesis*. Cham (CH): Springer.
- [9] Moreno, A. & Mossio, M. (2015). *Biological autonomy: a philosophical and theoretical enquiry*. Springer, New York.
- [10] Pormann, P. E. (2019). "Medical Conceptions of Health from Antiquity to Renaissance", in Adamson, P. (2019) (Ed.) *Health. A History*. Oxford: Oxford University Press, pp. 43-74.
- [11] Porter, R. (1998). *The Greatest Benefit of Mankind: A Medical History of Humanity from Antiquity to Present*. New York: Norton & Company.
- [12] World Health Organization. Regional Office for Europe (1984). Health promotion: a discussion document on the concept and principles: summary report of the Working Group on Concept and Principles of Health Promotion, Copenhagen, pp. 9-13 July 1984. Copenhagen, WHO Regional Office for Europe.

Reductionism and holism in pain medicine: dehumanization of medicine and complexity as the justification of clinical ethics

Alberto Monterde Fuertes*

Universidad del País Vasco/Euskal Herriko Unibertsitatea

The reductionism-holism debate is been a hot topic in medicine since Engel proposed the biopsychosocial model (Engel, 1977). The traditional reductionist stance considers that knowledge in one field can be translated to another (Andersen, 2016). This position is not suited for medicine and biology, because of the complexity of the phenomenon studied by those sciences (Andersen, 2016; Mitchell, 2009). Schaffner considers that the traditional reductionist position is characterized by greed reductions, but medical explanations are sweeping reductions: fragmented explanations that by themselves cannot account for the whole explanation of the disease (Schaffner, 2011). This explains why in medicine we find many compatible explanations for the same phenomena and why that plurality is necessary (Andersen, 2016; Mitchell, 2009).

However, explanatory reductionism is still a theoretical assumption in current biomedicine (Fuller, 2016). So, holists still critique biomedicine because of it. Holism is an elusive term, but is usually defined as “the whole is more than the sum of its parts” (Woods, 2017, p. 412). Some holistic authors aim to improve medical understanding and explanations and move them beyond reductionism (e.g: Anjum et al., 2020). Others consider that holism implies an ethical imperative in the patient’s treatment, so what ought to be corrected in medicine is the positivist epistemology of the clinical practice (Tauber, 2002). The later consider that problems in medicine are related to the loss of the human perspective (Marcum, 2008).

This debate started in the early stages of modern medicine and with the emergence of the physician as scientific expert (Rosenberg, 2007). Criticism against dehumanization had its origins in the introduction of technology and the rise of medical expertise, but was also liked to reductionism sometimes (Rosenberg, 2007). This resembles today’s arguments of humanist/holist medicine (e.g: Marcum, 2008). Similar criticism emerged in pain medicine in the past decade.

Some pain experts questioned the limitations of the traditional pain definition (Cohen et al., 2018; Williams & Craig, 2016), the reductionist assumptions in pain explanations

* amontf94@gmail.com

(Quintner et al., 2008; Brodal, 2017) and the influence that those ideas have in clinical praxis, treatment and the physician-patient relationship (Cohen et al., 2011; Quintner et al., 2016). They consider that reductionist perspective could help spreading wrong beliefs and attitudes regarding pain and promote the stigmatization of pain experience in chronic cases (Cohen et al., 2011; Quintner et al., 2016). So, their criticism is concerned with –and sometimes mimics– the threat of dehumanizing medical praxis, and argues that their causes are the inherent reductionist assumptions in medicine (Cohen et al., 2011; Quintner et al., 2016).

Those critics point towards something worth considering: the pragmatic implications of medical knowledge. This means that how we explain and conceptualize diseases has an impact upon medical praxis, treatment and maybe even patient-doctor relationship (Saborido, 2020). Thus, is worth exploring the explanatory and conceptual problems in pain medicine, and how reductionism can be misguiding clinical praxis.

Philosophers of science moved their explanatory perspective in medicine and biology from the nomological deductive model to mechanism because of the problems within the former (Lemoine, 2016). Mechanisms are defined as “a structure performing a function in virtue of its component parts, component operations and their organization” (Bechtel & Abrahamsen, 2005, p.423). They consist of entities and activities whose organization and interactions are responsible of the phenomena to explain (Illari & Williamson, 2012). In this sense, mechanisms causally explain the phenomena by describing the relationships that the different component parts have, the function that they realize, and how everything comes together to produce and orchestrated effect from which the phenomenon emerges (Bechtel & Abrahamsen, 2005).

The history of pain explanations is deeply rooted to the idea of pain as a sensation linked to tissue damage, and the consideration of that being sufficient to explain pain (Moayedí & Davis, 2013). Thus, pain was conceptualized as a uni-causal phenomenon. But recently, pain explanations shifted towards a multimodal approach, with multiple risk factors and variables that influence and modulate pain experience at different levels (Moayedí & Davis, 2013; Perl, 2007).

Such perspective shift calls for the introduction of the framework of complexity within explanations. Complexity allows us to understand why pain can't be explained by one mechanism, one causal factor or theory, and why it is modulated by different variables at different levels (from genetic to socioeconomic). The characteristics of complex biological systems like redundancy: a “built in fail-safe mechanism” (Mitchell, 2009), or robustness: “an outcome does not have a unique basis” (Mitchell, 2009), force us to understand pain as a non-linear phenomenon characterized by feedback loops, control and production mechanism (see: Bechtel, 2007).

If explaining pain is such a complex enterprise, is it right that reductionism could hinder patient-doctor relationship, treatment and the patient's experience of pain? Yes it is. Most treatments that are currently realized for low back pain are linked to a traditional uni-causal explanation of pain (see: Foster et al., 2018). Some of those treatments are not up to date with guidelines (Oliveira et al., 2018; Foster et al., 2018), and are related to the belief that pain is linked to tissue damage. Moreover, the nature of those treatments doesn't require a profound and close interaction with the patient or being aware of how the clinical practice affects her pain. Thus, accepting complexity in pain medicine means the introduction of a novel way of understanding and explaining pain, but it will also mean the justification of an ethical and up to date treatment: because how the patient is treated could influence pain and hinder recovery (Rossettini et al., 2018).

References

- [1] Andersen, H. (2016). Reduction in the biomedical sciences. En: Soomon, M., Simon, J. R., Kincaid, H. (eds.) *The routledge companion to philosophy of medicine*. London: Routledge.
- [2] Anjum, R. L., Copeland, S., Rocca, Elena (2020). *Rethinking Causality, Complexity and Evidence for the Unique Patient*. Springer.
- [3] Bechtel, W. (2007). *Mental mechanism: Philosophical perspectives on cognitive neuroscience*. New York: Routledge.
- [4] Bechtel, W., Abrahamsen, A. (2005). Explanation: A mechanist alternative. *Studies in the History and Philosophy of the Biological and Biomedical Sciences*, 36, 421-441.
- [5] Brodal, P. (2017). A neurobiologist's attempt to understand persistent pain. *Scandinavian Journal of Pain*, 15, 140-147.
- [6] Cohen, M., Quintner, J., Buchanan, D., Nielsen, M., Guy, L. (2011). Stigmatization of patients with chronic pain: the extinction of empathy, *Pain Medicine*, 12 (11), 1637-1643.
- [7] Cohen, M., Quintner, J., Van Rysewyk, S. (2018). Reconsidering the international association for the study of pain definition of pain. *Pain Reports*, 3 (2), e634.
- [8] Engel, G. (1977). The Need for a New Medical Model: A challenge for Biomedicine. *Science*, 196 (4286), 129-136.
- [9] Foster, E. N., Anema, R. J., Cherkin, D., Chou, R., Cohen, P. S., et al. (2018). Prevention and treatment of low back pain: evidence, challenges, and promising directions. *The Lancet*, 391 (10137), 2368-2383.

- [10] Fuller, J. (2016). *The new medical model: Chronic disease and evidence-based medicine* [Tesis de doctorado. Universidad de Toronto].
- [11] Illari, P. M., Williamson, J. (2012). What is a mechanism? Thinking about mechanisms across the sciences. *European Journal for Philosophy of Science*, 2 (1), 119–135.
- [12] Lemoine, M. (2016). Explanation in Medicine. In: Solomon, M. Simon, J. R., Kincaid, H. (eds). *The Routledge Companion to Philosophy of Medicine*. London: Routledge.
- [13] Marcum, J. (2008). *An introduction to philosophy of medicine: humanizing modern medicine*. Amsterdam: Springer Netherlands.
- [14] Mitchell, S. D. (2009). *Unsimple truths: science, complexity and policy*. Chicago: University of Chicago Press.
- [15] Oliveira, C., Maher, C., Pinto, R. Z., Traeger, A., Christine Lin, et al. (2018). Clinical practice guidelines for the management of non-specific low back pain in primary care: an updated overview. *European Spine Journal*, 11, 2791-2803.
- [16] Parris, W. C. V., Johnson, B. W. (2013). The history of pain medicine. En: Bazon, H., Rathmell, J. P., Wu, C. L., Turk, D., Argoff, C. E., Hurley, R. W. (eds) *Practical Management of Pain E-Book*. Elsevier Health Sciences.
- [17] Perl, E. R. (2007). Ideas about pain, a historical view. *Nat Rev Neurosci*, 8, 71-80.
- [18] Quintner, J. L., Cohen, M. L. (2016). The challenge of validating the experience of chronic pain: the importance of intersubjectivity and reframing. En Van Rysewyk, S. (ed.). *Meanings of pain*. Amsterdam: Springer, 281-293.
- [19] Quintner, J. L., Cohen, M. L., Buchanan, D., Katz, J. D., Williamson, O. D. (2008). Pain medicine and its models: helping or hindering? *Pain Medicine*, 9 (7), 824–834.
- [20] Rosenberg, C. (2007). Our Present Complaint. *American Medicine, Then and Now*. John Hopkins University Press.
- [21] Rossettini, G., Carlino, E., & Testa, M. (2018). Clinical relevance of contextual factors as triggers of placebo and nocebo effects in musculoskeletal pain. *BMC musculoskeletal disorders*, 19 (1), 27.
- [22] Saborido, C. (2020). *Filosofía de la Medicina*. Madrid: Tecnos.
- [23] Schaffner, K. F. (2011). Reduction in biology and medicine. *Philosophy of Medicine*, 16, 137.

- [24] Tauber, A. (2002). The Ethical Imperative of Holism in Medicine. En: Van Regenmortel, M. & Hull, D. (eds). *Promises and Limits of Reductionism in the Biomedical Sciences*. John Wiley & Sons.
- [25] Williams, A. C., Craig, K. D. (2016). Updating the definition of pain. *Pain*, 157 (11), 2420-2423.

Copies and dependence

Vittorio Morato*

Department of Philosophy, Sociology,
Education and Applied Psychology, University of Padua

Massimiliano Carrara**

Department of Philosophy, Sociology, Education and
Applied Psychology, University of Padua

According to N. Goodman (1976, p. 111), the relation of “being a copy of”, i.e. the relation between two distinct objects such as the second is a “reproduction” of the first, is “unexpectedly complex”. One could in effect be tempted to simply define the copying relation by means of similarity, treating “being a copy of” as a case of perfect similarity:

(C0) x is a copy of y iff x is perfectly similar to y .

A similarity approach to the relation of copy, however, is problematic. The problems are at least two. The first one is that of *accidental copies*: it could happen that two completely ontologically and historically independent objects end up being perfectly similar, without this implying that the first is a copy of the second. For example: imagine two perfectly similar diamonds discovered in two completely different places. The two are perfectly similar, but it would be strange, in such a situation, to claim that one is the copy of the other. For a start: which one of the two is the copy of the other? Which one is the original? The second problem is that of the *common original*: assume that, from an original object, two copies are made. The two copies are, by definition, perfectly similar to the original and thus perfectly similar between themselves. But it would be strange to say that the two copies are copies of one another. Again: which one is the copy? Which one is the original?

A simple way to fix the similarity approach is by adding intentionality to the recipe. One can say that:

(C1) x is a copy of y iff (i) x is perfectly similar to y and (ii) x is a product of intentional actions.

(C1) has the effect of limiting the domain of the relation to the products of (human) intentionality. There are no natural copies, according to C1. Notice, however, that a simple appeal to intentionality does not solve our problems: it is not difficult to imagine a scenario where two perfectly similar objects are the product of two separate and independent intentional actions and nonetheless are

* vittorio.morato@unipd.it

** massimiliano.carrara@unipd.it

not in the relation of copy with one another. Furthermore, (C1) is not even able to solve the problem of the common original: two copies of a common original are in fact the product of intentional actions, are perfectly similar to the original and thus between themselves, but, again, it would be strange to claim that one is the copy of the other.

A way to solve the problem of the common original is by making explicit that the copying relation gives rise to a form of *dependence* between a copy and its original:

(C2) x is a copy of y iff (i) y is perfectly similar to x , (ii) y is the product of intentional actions and (iii) y depends on x .

According to (C2), an object which is a copy is a copy of something else on which, somehow, depends. By means of (C2), two copies of a common original are not anymore in a relation of copying: two copies of a common original are not in any relation of dependence with one another. But what relation of dependence is relevant to characterise the copying relation? In current philosophical debate, the notion of dependence is hotly debated and discussed and there are various kinds of dependence relations available on the market: grounding, supervenience and various forms of ontological, conceptual, historic, counterfactual or causal forms of dependence (Tahko & Lowe, 2020). Aim of our paper is, first, to make explicit how elusive is the notion of dependence needed for the copying relation and, second, to capture what adequate kind of dependence for such a relation.

Let us consider, first, ontological dependence. In general, an object x is said to be ontologically dependent on another object y in case, necessarily, x exists only if y exists and it is not the case that, necessarily, y exists only if x exists (Lowe (1998); ?). Applying this definition to the relation of copying we will have the following:

(C3) x is a copy of y iff (i) y is perfectly similar to x , (ii) y is the product of intentional actions and (iii) necessarily, x cannot exist unless some original object y does and it is not the case that, necessarily, the original y exists only if x does.

The problem with this definition, however, is that a copy does not need the existence of its original to exist. The copied object, once it is created, is an ontologically independent object from the original. The original could go out of existence, but the copy could still continue to exist. Ontological dependence does not seem to be the right kind of dependence. One might try with weaker forms of dependence. Let us consider, for example, counterfactual dependence. In general, an object x counterfactually depends on y just in case had y not existed, x would not have existed (see Lewis (1973) for an application of counterfactual dependence to causation). Applying this kind of dependence to the copying relation, we would have the following condition:

(C4) x is a copy of y iff (i) y is perfectly similar to x , (ii) y is the product of intentional actions and (iii) had y not existed, x would not have existed.

But even in this case, we seem to have problems. The most important is that the relation of counterfactual dependence seems to overgeneralise: there are many objects over which an object is counterfactually dependent on, but these objects are not necessarily in the copying relation with such an object. This remains true even if we add the perfect similarity and the intentionality conditions. The same reasoning could be done for other kinds of dependence such as historic dependence (Thomasson, 2007).

These kinds of difficulties are made more serious if we consider some scenarios where the very same idea that the dependence relation needed to define the copying relation should be a rigid relation, namely a relation between specific objects. There are objects that are copies, but that are not copies of other specific objects. We will call this kind of copies non-objectual copies.

Consider some particular instances of copies which are counterfeits. A counterfeit is typically a copy of an original, specific, object, falsely presented as an original object, but there seem to be counterfeits that are copies but not copies of an object. Consider the notorious case of the forged Vermeers made by Van Meegeren. As laid out in most art history textbooks, Han van Meegeren was a Dutch painter who decided to prove his talent by forging paintings of some of the most famous artists, including Johannes Vermeer. He replicated so well the style and colours of the artist that the best critics and experts of the time regarded his paintings as genuine Vermeers. The forged Vermeers made by Van Meegeren, however, were not copies of some specific original painting made by Veermer.

If the existence of such non-objectual copies is accepted, all the forms of objectual dependence are out of the way. A way to emend the problem is by postulating some sort of sortal dependence:

(C5) x is a copy of y iff (i) y is perfectly similar to x , (ii) y is the product of intentional actions and (iii) x cannot exist unless an object y of a certain sort exists.

Even this condition, however, seems to be problematic. Consider the following mental experiment (due to Wreen (2002)): There are only sculptors. One day an artist, Ginus, announces to the world that the next day he would reveal a new form of art: painting. After a brief description of the nature of this new form of art, he isolates himself in his studio to produce the first painting but, during his solitude in the studio, he is kidnapped and a second artist, Minus, produces the first painting, making sure to attribute it to Ginus. If you accept this thought experiment, there is a counterfeit (i.e. a copy) but there isn't an object y of a certain sort whose x is a copy. This seems to be a non-sortal, non-objectual notion of copy.

In the final section of the paper, we will explore the idea that the copying relation has something to do with the essentiality of the origin:

(C6) x is a copy of y iff (i) y is perfectly similar to x , (ii) y is the product of intentional actions and (iii) necessarily, if an object y originates from an object x , then necessarily, if y exists, then y originates from x .

References

- [1] Goodman, N. (1976). *Languages of art* (2nd ed.). Indianapolis: Hackett.
- [2] Lewis, D. K. (1973). *Causation*. *Journal of Philosophy*, 70, 556-567. (Reprinted with a postscript in Lewis (1986, Ch. 21, pp. 159–213))
- [3] Lewis, D. K. (1986). *Philosophical papers* (Vol. II). Oxford: Oxford University Press.
- [4] Lowe, E. (1998). *The possibility of metaphysics*. Oxford: Clarendon Press.
- [5] Tahko, T., & Lowe, E. (2020). Ontological dependence. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy*. Retrieved from <https://plato.stanford.edu/archives/fall2020/entries/dependence-ontological/>>
- [6] Thomasson, A. L. (2007). *Ordinary objects*. Oxford: Oxford University Press. Wreen, M. (2002). *Forgery*. *Canadian Journal of Philosophy*, 32 , 143–166.

Disagreement and polarization in scientific communities

Javier Osorio*
Autonomous University of Madrid

The term “polarization” refers to several phenomena that share common features but are manifested in multiple ways (Bramson et al., 2017). The notions of polarization that deal with the socio-political dimension have been extensively studied for decades, and enjoy a broad theoretical and empirical support. A less researched phenomenon, but of great interest, is that of polarization in scientific communities. It is worth asking whether the dynamics of research in scientific environments can lead to the emergence of phenomena similar to those that have been detected by experimental psychology in contemporary societies.

A notion that allows us to understand certain types of dynamics in these scenarios is that of belief polarization. Belief polarization is a notion that accounts for those phenomena where, given a single body of shared evidence, two subjects tend to become increasingly divided towards their own prior beliefs (Pennebaker 2015, 94). Belief polarization is more easily manifested when subjects are in a group where values and beliefs are shared (Pomerantz, Chaiken, and Tordesillas, 1995) and when subjects have a well-developed cognitive structure (Chaiken and Yates, 1985; Liberman and Chaiken, 1991). One might think that, in this kind of phenomena, at least one of the two (or more) subjects in the dispute is committing some reasoning flaws. However, these subjects may be fully rational in acquiring and updating their beliefs. Several studies have been shown that the reasoning process that takes place in these scenarios is consistent with certain normative principles of revising and belief updating (see Nielsen and Stewart, 2021). Despite the fact that agents exhibit a fully rational information and updating processing, their beliefs are reinforced and polarized.

Belief polarization is closely related to a phenomenon that has been traditionally explored by epistemology and philosophy of science: disagreements between experts in scientific communities. This type of phenomenon has been investigated under the label of non-factual disagreement. A non-factual disagreement in scientific communities is a disagreement about the background, theoretical assumptions of the subjects, rather than being about some fact within the domain being investigated (Weinberger and Bradley 2020, 2). These assumptions influence, justify, and enable scientific practice; that is, they

* javiosorm@gmail.com

are an integral part of scientific activity (Andersen et al., 2019). On many occasions, a disagreement is of this nature in an indirect way, i.e., while the object of the disagreement is not explicitly a theoretical assumption or a set of them, the propositions that are indeed the object of the discussion depend derivatively on the assumptions on which they are based (Ranalli, 2018). This occurs because scientists are not always aware of the kind of theoretical assumptions they hold, since those assumptions operate inadvertently and are rarely made explicit (Rocca and Andersen, 2017).

This kind of phenomenon exhibits similar features to belief polarization. However, their relationship has not been made explicit in the literature. First, in both cases the evidence available to the subjects in dispute is not sufficient to resolve the disagreement, that is, the body of evidence alone does not determine a single answer toward the relevant set of facts. We are faced with cases of practical underdetermination of theory by evidence (Turnbull 2017, 6). Second, the divergence of opinions is partially determined by the influence of the assumptions or beliefs of the evaluating subjects. The “prior beliefs” of the experiments that measure belief polarization and the “theoretical assumptions” of subjects involved in non-factual disagreements are functionally identical elements, since the role they play is to guide the evaluation and assessment of the available evidence.

Richard Feldman has argued that these disagreements are not the product of agents’ rationality (Feldman, 2011). This line of inquiry by the epistemology of disagreement focuses exclusively on the rational character of subjects, but rarely attends to actual cases where the consequences can be highly detrimental. In this sense, the questions about the rational character of an individual’s action are independent of the questions about the practical consequences from the group’s joint action. Pace Feldman, research on the phenomenon of belief polarization suggests that disagreeing subjects are not necessarily being irrational. This, however, does not mean that the consequences that follow from such contexts are desirable. In other words: rationality is not, in these cases, guarantee of epistemic success.

When a situation requires the creation of policies to tackle a highly complex problem, this polarized scenario of experts offers two (or more) incompatible ways of positioning in the face of the same evidence. These contexts are unsuitable for generating scientifically informed policy-making plans where urgent decisions are required in areas that are imbued in uncertainty (Mason-Renton, 2019). So, to understand the dynamics and scope of such a controversy, we must not only look at whether the subjects are being sufficiently reasonable, we must also attend to the practical consequences that arise from such phenomena. For example, Evelyn Brister (2015) analyzes an interdisciplinary collaboration regarding environmental conservation policies in Central Africa. Disciplinary differences that have to do with the theoretical assumptions held by researchers give rise to a polarized debate among experts whose costs can be catastrophic: a poor performance

by the relevant agencies in charge of generating concrete policies can result in, on the one hand, the complete extinction of animal species; and on the other hand, contribute to the displacement, suffering and death of human beings and the fragmentation of population groups (Brister 2015, 89).

Detecting and analyzing these kinds of controversies where evidence is uncertain and expert opinions diverge can be helpful to better understand research and public decision-making processes (Mason-Renton, 2019). Making explicit the dynamics of these disagreements can be helpful in understanding those scientific controversies where the resolution of the disagreement by researchers and the action by those institutions that must make informed decisions is required. Experimental psychology shows that these kinds of scenarios, far from being a situation generated by bad reasoning or the systematic operation of cognitive biases, turn out to be a natural consequence of different ways of processing and evaluating the available evidence; philosophy of science shows that we should not only look at the rational character of group action, but also at the structure and practical consequences of these kinds of controversies at the interface of science and policy-making.

References

- [1] Andersen et al. (2019). *Philosophical bias is the one bias that science cannot avoid*. eLife, 8, 1-5.
- [2] Bramson et al. (2017). Understanding Polarization: Meanings, Measures, and Model Evaluation. *Philosophy of Science*, 84(1), 115-159.
- [3] Brister, E. (2016). Disciplinary capture and epistemological obstacles to interdisciplinary research: Lessons from central African conservation disputes. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 56, 82-91.
- [4] Chaiken, S. and Yates, S. (1985). Affective-cognitive consistency and thought-induced attitude polarization. *Journal of Personality and Social Psychology*, 49(6), 1470-1481.
- [5] Feldman, R. (2011). Reasonable religious disagreements. In A. Goldman y D. Whitcomb (Eds.), *Social epistemology: Essential readings* (pp. 137-157). Oxford and New York: Oxford University Press.
- [6] Kelly, T. (2005). Disagreement, Dogmatism, and Belief Polarization. *The Journal of Philosophy*, 105(10), 611-633.
- [7] Kuhn, D. and Lao, J. (1996). Effects of Evidence on Attitudes: Is Polarization the Norm? *Psychological Science*, 7(2), 115-120.

- [8] Liberman, A. and Chaiken, S. (1991). Value conflict and thought-induced attitude change. *Journal of Experimental Social Psychology*, 27(3), 203–216.
- [9] Mason-Renton et al. (2019). Science for Policy: A case study of scientific polarization, values, and the framing of risk and uncertainty. *Risk Analysis*, 39(6), 1229-1242.
- [10] Nielsen, M. and Stuart, R. T. (2021). Persistent disagreement and polarization in a Bayesian setting. *British journal for the philosophy of science*, 72 (1), 51-78.
- [11] Pennebaker, L. (2015). The Epistemological Significance and Implications of Belief Polarization. *Res cogitans*, 6(1), 93-101.
- [12] Pomerantz, E. M., Chaiken, S., and Tordesillas, R. S. (1995). Attitude strength and resistance processes. *Journal of Personality and Social Psychology*, 69, 408-419.
- [13] Ranalli, C. (2018). What is Deep Disagreement? *Topoi*, 1(3), 1-16.
- [14] Rocca, E. y Andersen, F. (2017). How biological background assumptions influence scientific risk evaluation of stacked genetically modified plants: an analysis of research hypotheses and argumentations. *Life Sciences, Society and Policy*, 13(11), 1-20.
- [15] Turnbull, G. M. (2017). Underdetermination in science: What it is and why we should care. *Philosophy Compass*, 13(2), 1-11.
- [16] Weinberger, N. y Bradley, S. (2020). Making sense of non-factual disagreement in science. *Studies in History and Philosophy of Science Part A*, 83, 36-43.

There are no “genuine mathematical explanations” in biology, but mathematical heuristics

José Antonio Pérez Escobar*
ETH Zurich

In his 2005 paper, Alan Baker proposed an example of “genuine mathematical explanation” in biology: the cicada life-cycle being prime-numbered, because a prime-numbered life-cycle minimizes the intersection with their predator’s life-cycle. Lyon and Colyvan (2008) discuss another paradigmatic case in biology: the hexagonal shape of honeycomb cells. Because bees are evolutionarily pressured to use as little wax as possible in building the perimeter of the cells, and these cells must be able to store as much honey as possible. Both Baker, and Lyon and Colyvan, identify that there are both biological and mathematical parts in the explanation of the cicada and honeycomb cases, respectively. My contention is that the biological character of those parts where mathematics intervenes is not recognized. Baker’s fifth explanandum, the prime-numbered-year cicada life-cycle, is according to him the one to be explained mathematically (others, like “the great duration of the cicada life-cycle”, are to be explained biologically). Lyon and Colyvan do something similar: bees not wasting wax is to be explained biologically, but the honeycomb cells being shaped hexagonally is to be explained mathematically. This distinction is contrived and blind to the specificities of biological discursive entities.

Biology is characterized by employing and requiring teleological/functional explanations (Ayala 1968; Nagel 1977; Mayr 1992; Ayala 1999). Broadly conceived, these explanations consist in making sense of the existence and/or the characteristics of the explanandum by appealing to its purposes/functions¹.

While teleological/functional explanations are heterogeneous, there have been many attempts to describe their general structure. One of the few formalizations of the general structure of teleological/functional explanations is provided by Wimsatt (1972). Very simply, Wimsatt’s account can be read as: given causal laws T , the function of the behavior B of a given item i in system S in environment E relative to purpose P is to do C . P has explanatory power over $B(i)$: it explains its characteristics, existence, or both. Note that P

* jose.perez@gess.ethz.ch

¹ I will not be concerned here with the metaphysical implications of the use of this type of explanations (do they imply teleology in a metaphysically strong sense? Or are they mere heuristics in the description of a mechanical world?). Therefore, I will not distinguish here between terms like “teleology”, “teleonomy” or “function”, since such a distinction is relevant for the metaphysical debate, but not so much for the logical structure of explanations in biology.

is not reducible to (B(i), S, E) and C, but it is underdetermined by the physical description of the set. P can, however, be inferred by certain means (asking the designer of a given artifact, or guessing the evolutionary story of B(i) in S and E), and then ascribed to B(i).

Wimsatt's formula provides a framework to understand the hexagonal shape of honeycomb cells in a way more akin to biological understanding:

B(i): hexagonal cells in the honeycomb

S: The honeycomb and bee colony

E: the ecosystem where S belongs

Cs: Honey stored and wax used, as a result of a certain cell shape

P1: Storing as much honey as possible

P2: Using the least amount of wax possible

According to the formula, P explains B(i). In this case, the fact that the P of the honeycomb cells is to store as much honey as possible, while requiring the least amount of wax possible, explains the hexagonal shape of the honeycomb cells. It turns out that the hexagonal arrangement is an efficient solution to fulfill a certain compromise between P1 and P2. But is most of the explanatory burden carried by mathematics? There are at least two reasons why it is not, but it is carried by P instead.

1) Certain B(i)s are optimally functional only relative to the relevant Ps at stake. The hexagonal perimeter shape would not be the best shape if only P2 applied; the best shape would be a circular shape instead, or not having individual cells at all. However, P1 and P2 interfere with each other, and a solution to satisfy both is a hexagonal cell shape. Similarly, if another relevant P, P3, applied, and this interfered with P1 and P2, perhaps a different solution would be better suited to fit them all. A likely P3 is structural integrity, but this P likely presents no interference.

2) Certain B(i)s are optimally functional only relative to the comparative importance of the different Ps at stake. Hexagonal cells constitute the most optimal way to fill all available space, employing the least amount of wall perimeter in doing so. That is, P1 has a considerable primacy over P2. However, if P2 had primacy over P1 (if wax was a more scarce resource than it already is, or honey was less valuable or necessary in smaller amounts, or space more abundant), the most optimal solution may be, again, circular shapes, of different sizes depending on the relative weight of each of the Ps. While optimal solutions, informed by mathematics, depends on P ("optimal for what?"), the opposite does not hold true. P is a more fundamental element in the explanation.

At the base of the reasons why mathematics does not carry the explanatory burden is a distinction between mathematical optimization and biological optimization. In the

context of biological optimization, mathematics is a heuristic for a given evolutionary trajectory. Let us consider the main points of divergence:

- 1) Mathematical optimization requires a quantitative measure of “cost” to be specified. “X is the mathematically optimal way to do Y” assumes such a quantitative measure. However, this is not granted in biology, where different, contingent and dynamic combinations and weightings of resources are biological “cost”.
- 2) The choice of cost function is contingent on scope too. Cost optimization for a given P alone may be irrelevant if we factor in other Ps, and again, P configurations are dynamic, variable. Evolutionary trajectories can follow different paths of optimization yielding different results.
- 3) The basic tenet of a mathematical explanation is that, among all relevant mathematical objects, X is the best (or necessary) in terms of measure Y. This, however, does not apply to biology, for two reasons. First, all mathematical objects relevant for the cost function is a much bigger set of objects than those biologically feasible. Second, evolution does not take into account all biologically feasible solutions and chooses the best for the cost function. Instead, it pits different locally emergent solutions against each other. Therefore, a resulting biological phenomenon is not explained in that it is the best of all mathematical solution, but in that it was the best biological option of those readily locally available.

Hence, not only the shape of the honeycomb cells is better explained by P configurations rather than mathematics, but also by the character of biological optimization and evolutionary pressure, which differs from mathematical optimization. The fact that in some cases there may be a phenomenological overlap with mathematical objects should not mask the distinction. These considerations, in line with the specificities of biology as a field, entail a nominalistic reconstruction of the explanation which does not point at a mathematical object. Here, the mathematical insight plays two parts:

- 1) It assists P discovery (or P ascription). Sometimes, “clean mathematical” features may arise when B(i) does not have many complicated Ps of similar importance substantially interfering with each other. When the specific Ps of a biological object are unknown, “clean mathematical” features may stand out and help figure them out. For instance, circular shapes usually indicate that surface material (like membranes, or shells) is expensive in terms of biological cost, and volume must be maximized, but outer space is not a concern. To the degree that space is a concern, like in bird’s oviducts, then a compromise on the round shapes due to an interfering P may arise, in the form of oval shapes².

Because round shapes, more common in biology, are not as striking as hexagons, they have not been given the same explanatory treatment. This realization trivializes the whole issue, as mathematical explanations would subsume much of biology.

² Further compromises lead to subsequent departures from “clean mathematical” features, explaining the rest of biological shapes.

2) It bridges P and B(i), enabling teleological/functional explanations. In the case of the hexagonal shape of honeycomb cells, mathematics (in this case, basic geometry) may help heuristically understand why a certain solution (here, hexagonal shapes) is an efficient fulfiller of a given P configuration (space and wax economy). Thus mathematics enable the explanation of B(i) in terms of P by adding an extra mediating insight.

References

- [1] Baker, A. (2005). Are there Genuine Mathematical Explanations of Physical Phenomena?. *Mind*, 114, 223-238.
- [2] Lyon, A. & Colyvan, M. (2008). The Explanatory Power of Phase Spaces. *Philosophia Mathematica*, 16, 1-17.
- [3] Mayr, E. (1992). The idea of teleology. *Journal of the History of Ideas*, 53 (1), 117-135.
- [4] Nagel, E. (1977). Teleology revisited: goal-directed processes in biology. *Journal of Philosophy*, 74, 261-301.
- [5] Wimsatt, W. C. (1972). Teleology and the Logical Structure of Function Statements. *Studies in History and Philosophy of Science* 3 (1-80).

Mechanisms and policymaking: lessons from EBM+

Saúl Pérez-González*
University of Turin

The evidence-based policy (EBP) movement, inspired by evidence-based medicine (EBM), argues that policy development, evaluation, and implementation should be guided by the best available evidence. Available evidence about policy interventions is assessed on the bases of evidence hierarchies, which focus on evidence-gathering methods and their potential to suffer from systematic biases (Cartwright & Hardie, 2012). According to those hierarchies, the highest quality evidence is provided by randomized control trials (RCTs) and meta-analysis of RCTs. Mechanical reasoning and case reports are usually placed on the bottom of the hierarchies. In the last decades, EBP has reached great relevance in areas such as public health and development economics. This relevance became particularly clear with the awarding in 2019 of the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel to Abhijit Banerjee, Esther Duflo, and Michael Kremer.

Nonetheless, in the last two decades, evidence hierarchies and RCTs-based approaches have been subject to discussion (Cartwright, 2013; Stegenga, 2014; Victora et al., 2004; Worrall, 2002). One of the main concerns is that causal inferences based on evidence of difference-making provided by RCTs face important difficulties. For instance, they can hardly deal with confounding factors, non-causal correlations, and causal extrapolations from one population to another. Given its deficiencies, it has been argued that evidence of difference-making should be complemented by evidence of mechanisms (Clarke et al., 2013, 2014; Russo & Williamson, 2007). Advocates of evidential pluralism consider that evidence of mechanisms addresses the main weaknesses of evidence of difference-making (e.g., confounding factors), and vice versa. Evidence of mechanisms is understood as evidence of the existence or the properties of mechanisms in the domain of inquiry (Illari, 2011). In that context, a mechanism can be a complex system, a causal process, or some combination of both.

The pluralist approach has mainly focused on EBM, leading to the development of EBM+ (Clarke et al., 2013, 2014; Parkkinen et al., 2018). Within that pluralistic

* saul.pergon@gmail.com

framework, relevant issues such as drug development and approval (Aronson et al., 2018), risk assessment (Rocca, 2017), and clinical practice (Tonelli & Williamson, 2020) have been addressed. Nevertheless, much less attention has been given to evidential pluralism in EBP. The aim of this paper is to, on the bases of the previous debate on EBM+, discuss the prospects of evidential pluralism in EBP. For that purpose, three aspects will be taken into account:

1. *What relevant (or problematic) aspects could be addressed by evidence of mechanisms?* The main argument in support of EBM+ is that evidence of mechanisms assists us in important areas such as identification of effective treatments and extrapolation from animal models to humans. In order to assess evidential pluralism in EBP, it is crucial to specify how could it contribute to policy development, evaluation, and implementation.
2. *Would problematic aspects of evidence of mechanisms considerably undermine their contribution?* Critics of EBM+ have point out that evidence of mechanisms faces important difficulties such as masking, fragmentary knowledge, and absence of regularity (Howick, 2011; Howick et al., 2013). Consequently, to actually assess the potential of evidential pluralism, the relevance of those difficulties in social science and policymaking should be considered.
3. *How real scientific practice engages with the pluralistic framework?* Advocates of EBM+ appeal to historical (e.g., Bradford Hill) and contemporary (e.g., International Agency for Research on Cancer (IARC)) biomedical practice to support the complementarity between evidence of difference-making and evidence of mechanisms. To analyse the possible contribution of evidence of mechanisms in EBP, it is important to consider the roles it has (or could have) played in policymaking and evaluate its performance.

In order to address those questions and asses evidential pluralism in EBP, special attention will be paid to heterogeneity of mechanisms. New mechanists consider that mechanisms are relevant in many scientific fields and aim to develop general mechanical approaches that encompass all of them (Glennan, 2017; Illari & Williamson, 2012). Nevertheless, mechanisms may differ a lot from one field to another and general approaches face important difficulties (Author, 2019). Consequently, although the previous debate on evidential pluralism in EBM constitute a valuable resource for assessing its relevance in EBP, particularities of social mechanisms and evidence of social mechanisms must be taken into account. It cannot be straightforwardly concluded that evidential pluralism would have the same strengths and weaknesses in EBP.

References

- [1] Author. (2019).
- [2] Aronson, J. K., Caze, A. L., Kelly, M. P., Parkkinen, V.-P., & Williamson, J. (2018). The use of mechanistic evidence in drug approval. *Journal of Evaluation in Clinical Practice*, 24(5), 1166-1176.
- [3] Cartwright, N. (2013). Knowing what we are talking about: why evidence doesn't always travel. *Evidence & Policy: A Journal of Research, Debate and Practice*, 9(1), 97-112.
- [4] Cartwright, N., & Hardie, J. (2012). *Evidence-based policy: A practical guide to doing it better*. Oxford University Press.
- [5] Clarke, B., Gillies, D., Illari, P., Russo, F., & Williamson, J. (2013). The evidence that evidence-based medicine omits. *Preventive Medicine*, 57(6), 745-747.
- [6] Clarke, B., Gillies, D., Illari, P., Russo, F., & Williamson, J. (2014). Mechanisms and the evidence hierarchy. *Topoi*, 33(2), 339-360.
- [7] Glennan, S. (2017). *The new mechanical philosophy*. Oxford University Press.
- [8] Howick, J. (2011). Exposing the Vanities –and a Qualified Defense– of Mechanistic Reasoning in Health Care Decision Making. *Philosophy of Science*, 78(5), 926- 940.
- [9] Howick, J., Glasziou, P., & Aronson, J. K. (2013). Problems with using mechanisms to solve the problem of extrapolation. *Theoretical Medicine and Bioethics*, 34(4), 275-291.
- [10] Illari, P. M. (2011). Mechanistic evidence: disambiguating the Russo-Williamson thesis. *International Studies in the Philosophy of Science*, 25(2), 139-157.
- [11] Illari, P. M., & Williamson, J. (2012). What is a mechanism? Thinking about mechanisms across the sciences. *European Journal for Philosophy of Science*, 2(1), 119-135.
- [12] Parkkinen, V.-P., Wallmann, C., Wilde, M., Clarke, B., Illari, P., Kelly, M. P., Norell, C., Russo, F., Shaw, B., & Williamson, J. (2018). *Evaluating evidence of mechanisms in medicine: principles and procedures*. Springer.
- [13] Rocca, E. (2017). Bridging the boundaries between scientists and clinicians— mechanistic hypotheses and patient stories in risk assessment of drugs. *Journal of Evaluation in Clinical Practice*, 23(1), 114-120.

- [14] Russo, F., & Williamson, J. (2007). Interpreting causality in the health sciences. *International Studies in the Philosophy of Science*, 21(2), 157-170.
- [15] Stegenga, J. (2014). Down with the Hierarchies. *Topoi*, 33(2), 313-322.
- [16] Tonelli, M. R., & Williamson, J. (2020). Mechanisms in clinical practice: use and justification. *Medicine, Health Care and Philosophy*, 23(1), 115-124.
- [17] Victora, C. G., Habicht, J.-P., & Bryce, J. (2004). Evidence-based public health: moving beyond randomized trials. *American Journal of Public Health*, 94(3), 400-405.
- [18] Worrall, J. (2002). What Evidence in Evidence-Based Medicine? *Philosophy of Science*, 69(S3), S316-S330.

Philosophical Foundations for an Empirical Bayesianism Interpretation of Null Hypothesis Significance Testing in Psychology

Eoin Perry*
University of Bristol

Bird (2018) suggests that the perceived ‘replication crisis’ in psychology is to be expected; psychological hypotheses have a sufficiently low prior-probability that their posterior- probability remains relatively low, given that they are found significant in a Null Hypothesis Significance Test (NHST). This kind of use of prior probabilities to interpret NHST studies has been argued to contradict and misapply the NHST statistical paradigm. Such arguments are given by Mayo, whose ‘severe testing’ philosophy (see Mayo 1996, 2018) offers the best way to understand and defend the NHST paradigm. By examining how Bayesian interpretation of NHSTs can be justified within an amended version of Mayo’s framework, this paper; (1) shows how Severity analysis can function as a foundational evidentiary principle, justifying the particular probability assignments (‘Bayesian-Probabilities’) used in Bayesian interpretations of psychological research, (2) illustrates a potentially useful methodological innovation involving Bayesian interpretation of NHSTs, (3) clarifies the scope of applicability of some cogent objections, due to Mayo and Morey, to Bayesian interpretation of NHSTs.

I briefly outline an epistemological framework that motivates, and makes epistemologically coherent, the use of Bayesian models with severely tested Bayesian-Probabilities. Following Spanos & Mayo (2015), Cartwright (1997), in this framework, probability distributions over outcomes only arise if some real ‘mechanism’ generates these distributions. The framework follows Pettigrew (2013a) in holding that the degrees of *credence* we assign to statements should minimize the sum of the squared deviations of those credences from the statements’ truth values. As I discuss, to achieve this, credences should; (1) respect Kolmogorov’s probability axioms (Pettigrew, 2013a), and (2) roughly speaking be as close as possible to hypotheses objective probabilities (Pettigrew; 2013b). The role of severe testing in this framework is to provide a rationale for assigning credences to hypotheses, proportional to the degree of evidence for those hypotheses, when they have no prior or posterior probabilities, as may occur when we wish to determine which of a set of possible mechanisms might be generating some data.

* eoin.perry07@gmail.com

I show how the apparently contradicting interpretations of ‘evidence’ involved in the Bayesian and Severity approaches to statistical inference can be reconciled and combined. In contrast to Bayesian views, severity approaches do not identify the evidence for a hypothesis, H , with a probability assigned to H , but rather with how severely H has been tested. One thing involved in making this Severity assessment is Mayo’s ‘Frequentist Principle of Evidence (FEV)’. Paraphrasing Mayo (e.g. Mayo & Spanos 2006 p. 82), this can be presented as follows:

Negative FEV

*Data, d , provide good evidence **against** H_0 to the extent that, were d generated by the mechanism implied by H_0 , then it is unlikely that we would have observed data as far or further than d from their expected value under H_0 .*

Positive FEV

*If H_a is logically or mathematically implied by the rejection of H_0 , then data, d , provide severe evidence **for** H_a to the extent that d provide severe evidence against H_0 .*

Howson’s (1997) Bayesian critique of Severity Theories can be framed as attributing to Severity-Theorists, through their advocacy of FEV, the denial of what I call ‘The Comparative Probability Principle’ (CPP):

If our evidence indicates that H_1 is more probable than H_2 , we have better evidence for H_1 than H_2 .

I rebut Mayo’s (1997, p. 209; 2018; p. 368) arguments which seem to imply a rejection of CPP and reject as insufficiently clarificatory her suggestion that Howson misapplies FEV. I then provide a response to Howson, similar to that of Spanos (2010), which involves amending FEV to accommodate CPP. Here is an amended FEV:

Negative FEV*

*Data, d , provide good evidence **against** H_0 to the extent that were d generated by the mechanism implied by H_0 , then it is unlikely that we would have observed data as far or further than d from their expected value under H_0 .*

So long as...

H_0 describes the most general mechanism which might generate d , which d provide good evidence against as per (1).

I show how FEV* allows the Severity-Theorist to maintain that the evidential support for a set of hypotheses is closely proportional to the probabilities assigned to them by a severely tested Bayesian model, with the severity of the evidence for that model providing

a meaningful estimate of the higher-order uncertainty around these probabilities. And, as I show, since the *model* is severely tested, we can defuse Mayo's (2010, p. 197) objection that this would allow us to apportion 'evidence' to hypotheses in a way that fails to meet some of the desiderata of severe testing.

I articulate an epistemic rationale for FEV*, referencing Mayo's (1996 p. 95) arguments for the original FEV, and Sprenger's and Spielman's (2016, 1974) discussions of Fisher's interpretation of NHSTs. Suppose we must assign credences to a set of mathematically exhaustive hypotheses, $[H_1 \dots H_n]$, based on some data, d . Suppose that $[H_1 \dots H_n]$ each have no prior probability but that d gives each member of $[H_1 \dots H_n]$ a different likelihood (i.e. for each member, H , of $[H_1 \dots H_n]$ d has some probability conditional on H). If we assign, say, a 0.1 credence to a member, H , of $[H_1 \dots H_n]$ whose likelihood is 0.1, this amounts to assigning 0.1 credence to the claim 'H is true and the data turned out as they did while having had an objective probability to do so of 0.1'. Thus, we can adapt Pettigrew's (2013b) argument for assigning credences to propositions in line with their objective chances to suggest that credences should be assigned to $[H_1 \dots H_n]$ in line with their likelihoods given d .

As I will show, when applied to equivalence tests (a type of NHST described by Lakens 2017), severity reasoning recommends accepting an estimate for the value of a parameter which is an interval around the value recommended by likelihood analysis, the width of which is determined by one's preferred evidential threshold for accepting hypotheses. This is compatible with assigning credences in the way recommended by the above Likelihoodist analysis. If the argument for severity is that we can reach the same inferences as the Likelihoodist, why not just be a Likelihoodist? One reason, demonstrated by Mayo & Kruse (2001), which I will briefly discuss, is that the Likelihoodist cannot resolve the 'optional stopping problem' whereas the Severity-Theorist can. Thus we can argue that severity analysis is necessary to justify the particular Bayesian-Probabilities used in Bayesian models.

Referencing literature on p-curve analysis (Simonsohn et al. 2013), 'super-forecasters' (Mellers et al., 2017), and the benefits of precise probabilistic forecasts for political events (Friedman et al., 2017), I consider the prospects for using severe tests to establish that hypotheses from particular individual psychologists, or sub-fields, have certain Bayesian- Probabilities. This idea closely resembles the 'empirical Bayesian' interpretation of NHSTs which is advocated for psychology by Schimmack (2021), and has found recent application in genomics (Bickel, 2020).

P-curves are distributions of p-values amongst sets of studies. As Simonsohn et al. (2013) show, a given average effect size amongst a set of studies $[S_1 \dots S_n]$ implies an expected distribution of p-values. I show how examining a p-curve for $[S_1 \dots S_n]$ can provide severe evidence, as per FEV, that the average effect size amongst $[S_1 \dots S_n]$ is no greater

than some threshold (or no lower if we can rule out the presence of substantial p-hacking based on information provided in pre-registration reports). If for some $[S_1 \dots S_n]$ we establish with severe evidence that $[S_1 \dots S_n]$ must have an average effect size no greater than 0.2, and we encounter a further (proposed) study, S^* , which appears relevantly similar to $[S_1 \dots S_n]$ then we can infer that the true prior effect size for S^* is no greater than 0.2. We can then obtain a posterior expected effect size based on the p-value obtained in S^* . I discuss how the inductive nature of the inference that S^* relevantly resembles $[S_1 \dots S_n]$ introduces unknown bias and Knightian uncertainty into the obtained posterior effect size estimate for a hypothesis. Referencing Rubin's (2018) analysis of NHSTs in psychology, I argue that analogous uncertainty is involved in the standard epistemic interpretation of NHSTs adopted by almost all psychological studies which use them.

I assess two anticipated objections (drawn from Mayo and Morey, 2018) to this proposed interpretation of NHSTs. According to these arguments, the Bayesian-Probabilities to be used in such analyses are; (1) subject to reference class uncertainty, and (2) derived in a way that involves the 'probabilistic instantiation fallacy'. I consider how we might mitigate these problems by eliciting additional information from psychologists in pre-registration reports (of the kind advocated by Lakens 2019). For example, we can mitigate the reference class uncertainty problem involved in attributing a prior expected effect size to a study by psychologist A by grouping it with other studies by A, about similar topics, that were given similar estimated prior expected effect sizes by A, in pre-registration reports.

Referencing Meehl's (1992) discussion of how we might empirically scrutinize different scientific methodologies, I argue that a justification for this methodological innovation would need to be pragmatic and empirical, perhaps involving further higher-order severity analyses.

References

- [1] David R. Bickel (2020). 'Genomics Data Analysis: False Discovery Rates and Empirical Bayesian Analysis, *CRC Press*.
- [2] Bird; Alexander (2018). 'Understanding the Replication Crisis as a Base Rate Fallacy', *The British Journal for the Philosophy of Science*. DOI: <https://doi.org/10.1093/bjps/axy051>
- [3] Cartwright; Nancy (1997). 'Where do Laws of Nature Come From', *Dialectica*, 51(1), pp. 65-78. DOI: <http://www.jstor.org/stable/42970706>
- [4] Friedman; Jeffrey A., Baker; Joshua D., Mellers; Barbara A., Tetlock; Philip E., Zechhauser; Richard (2017). 'The Value of Precision in Probability Assessments: Evi-

- dence from a Large Scale Geopolitical Forecasting Tournament, *International Studies Quarterly*, 62 (2), pp 1 -13. DOI: <https://doi.org/10.1093/isq/sqx078>
- [5] Howson; Colin (1997) 'Error Probabilities in Error', *Philosophy of Science*, 64. Stable URL: <https://www.jstor.org/stable/188402>
- [6] Mayo; Deborah (1996). 'Error and the Growth of Experimental Knowledge', *The University of Chicago Press*.
- [7] Lakens; Daniel (2017). 'Equivalence Tests: A Practical Primer for t-tests, Correlations, and Meta-Analyses. *Social Psychological and Personality Science*, 8(4), 355-362. DOI: 10.1177/1948550617697177
- [8] Lakens; Daniel (2019). 'The Value of Pre-registration for Psychological Science: A Conceptual Analysis', *Japanese Psychological Review*, 62(3), 221-230. DOI: https://doi.org/10.24602/sjpr.62.3_221
- [9] Mayo; Deborah (1996). 'Error and the Growth of Experimental Knowledge', *The University of Chicago Press*.
- [10] Mayo; Deborah (1997). 'Error Statistics and Learning from Error: Making a Virtue of Necessity, *Psychological Science*. DOI: 10.1086/392600
- [11] Mayo; Deborah (2018). 'Statistical Inference as Severe Testing: How to Get Beyond the Statistics Wars', *Cambridge University Press*.
- [12] Mayo; Deborah & Morey; Richard (2018). 'A Poor Prognosis for the Diagnostic Screening Critique of Statistical Tests', *OSF Preprints*. DOI: 10.31219/osf.io/ps38b
- [13] Mayo; Deborah & Spanos; Aris (2006). 'Severe Testing as a Basic Concept in a Neyman Pearson Philosophy of Induction', *The British Journal for the Philosophy of Science*, 57(2). DOI: <https://doi.org/10.1093/bjps/axl003>
- [14] Mellers; Barbara A., Metcalfe Bishop; Michael, Murray; Terry, Chen; Eva (2015) 'Identifying and Cultivating Super forecasters as a Method of Improving Probabilistic Predictions', *Perspective on Psychological Science* 10(3), pp. 267-281. DOI: <https://doi.org/10.1177%2F1745691615577794>
- [15] Meehl; Paul (1992). Cliometric Meta-Theory: The Actuarial Approach to Empirical History- Based Philosophy of Science, *Psychological Reports*, 71(2), 339–467.
- [16] Pettigrew; Richard (2013a). 'Epistemic Utility and Norms for Credence', *Philosophy Compass*, 8/10. DOI: 10.1111/phc3.12079
- [17] Pettigrew; Richard (2013b). 'A New Epistemic Utility Argument for the Principal Principle', *Episteme*, 10, pp 1935. DOI: 10.1017/epi.2013.5

- [18] Rubin; Mark (2018). 'What Type of Type-1 Error, Contrasting the Neyman Pearson and Fisherian Approaches in the Context of Exact and Direct Replications', *Synthese*, DOI: <https://doi.org/10.1007/s11229-019-02433-0>
- [19] Simonsohn; Uri, Nelson; Leif D., Simmons; Joseph P., (2013). DOI: 10.1037/a0033242
- [20] Schimmack; Ulrich (2021). 'Men are Created Equal: P-Values Are Not', blog post. Available at: <https://replicationindex.com/2021/01/15/men-are-created-equal-p-values-are-not/>
- [21] Spanos; Aris (2010). 'Is Frequentist Testing Vulnerable to the Base Rate Fallacy', *Philosophy of Science* 77(4). DOI: <https://doi.org/10.1086/656009>
- [22] Spanos; Aris & Mayo; Deborah (2015). 'Error Statistical Modelling and Inference: Where Methodology Meets Ontology', *Synthese*, 192:3533–3555. DOI: 10.1007/s11229-015-0744-y
- [23] Spielman; Stephen (1974). 'The Logic of Tests of Significance', *Philosophy of Science*, 41(3), pp. 211-226. Stable URL: <http://www.jstor.org/stable/187132>
- [24] Sprenger; Jens (2016). 'Bayesianism vs Frequentism in Statistical Inference', in: 'The Oxford Handbook of Probability and Philosophy', e.d. Hájek; Alan & Hitchcock; Christopher, *Oxford University Press*.

Theoretical innovation and scientific explanation. Two intertwined tasks

Andrés Rivadulla*
H. Prof., Complutense University of Madrid

1 Introduction

This article focuses on the relationship between innovation and explanation, two fundamental tasks of theoretical science. In the examples of theoretical explanation usually offered the explanation takes place in the context of a single theory, i.e. the explanation is *intra-theoretical*. Nonetheless there is a major component of science to which contemporary philosophy has not given the attention it deserves. I call it *theoretical innovation* and it consists in devising unexpected, sometimes surprising, and in any case useful proposals for the advancement of science. Theoretical innovation is a form of scientific *creativity* and arises very often through the deductive discovery practice that I call *preduction*. As I have been claiming in Rivadulla 2008, 2010b, and elsewhere, what characterizes preduction as a deductive form of scientific discovery is that it involves, as premises, accepted results from several theories and/or disciplines. It is, therefore, an *inter-theoretical* way of reasoning. The important thing about preduction is that it also serves the purposes of theoretical explanation. Thus scientific explanation based on preduction is a form of inter-theoretical explanation as well.

It is truly amazing that, in the years when the so-called *context of scientific discovery* was banned as methodologically irrelevant by philosophers in the second half of the thirties of the last century, Louis De Broglie (1892-1987) devoted chapter 5, “On invention in theoretical sciences” of his 1941 book, to scientific discovery. For De Broglie (1941, *op. cit.*) the theoretical invention was in some way a discovery: “both, the creation and the invention in the domain of theoretical science possesses in certain aspects the characters of the discovery”. Theoretical inventor (*théoricien inventeur*) and innovative theorist (*théoricien novateur*) are two expressions that De Broglie used interchangeably.

* arivadulla@filos.ucm.es

2 Innovation and intra-theoretical explanation

Carl Hempel (1965, p. 345) made use already of the expression ‘theoretical explanation’. In Rivadulla (2005, pp. 169-177; 2010a, pp.12-14 and 2019 §18.8) I affirm that every physical construct, be it a fact, an empirical generalization, an abstract law and even a theory itself, receives a theoretical explanation when it is deduced mathematically within the framework of a broader physical construct. As an illustration I present below the theoretical explanation, within the framework of special relativity, of the time dilation in the decay of muons.

In what is known as the *zoo* of elementary particles, physicists distinguish between *Hadrons* and *Leptons*. *Hadrons* are particles subjected to nuclear strong interaction and include as sub-kinds, baryons and mesons. Protons and neutrons are baryons, particles composed by three quarks. Mesons are particles consisting of one quark and one anti-quark, like pions and kaons

For their part *leptons* are particles which interact only through electromagnetic and weak interaction. The electron, the tauon, the muon and the three types of neutrinos associated with them are leptons. In particular, muons, μ , are 207 times more massive than electrons and they have a mean life $\tau = 2.2 \times 10^{-6}$ sec, i.e. they are unstable, giving rise to an electron, an electron antineutrino, and a muon neutrino in their decay.

Well, since muon decay is a case of particle disintegration, we only need to resort to nuclear physics to calculate the number of radioactive nuclei that remain from an initial sample, after a given time (Rivadulla 2004, pp. 62- 63). This is the model that Ostlie & Carroll (1996, pp. 105-106) follow.

As a matter of fact, if $N(t)$ designates the number of muons remaining at time t from an initial sample, the number of muons that decay in a fraction dt of τ will be

$$dN(t) = -N(t) \frac{dt}{\tau}.$$

Integrating between $N(0)$ – the number of muons in the initial sample – and $N(t)$ on the left and between 0 and t on the right; making use of the operation of difference of logarithms and applying the definition of natural logarithm,

$$N(t) = N(0)e^{-t/\tau}$$

gives the number of remaining muons after a time t .

According to David Frisch and James Smith (1963, p. 343),

μ -mesons ... are produced high in the atmosphere and come shooting down toward the Earth with a speed greater than $0.99c$. As they come down, some of

them disintegrate in flight. The number arriving at medium altitude is, therefore, greater, than the number surviving to reach sea level.

To verify this, Frisch & Smith placed a muon detector at the top of Mount Washington, New Hampshire, which counted 563 muons/hour which fell 1907m down the mountain to sea level with a velocity $v=0.9952c$ in $\frac{1907m}{0.9952c} = 6.4 \times 10^{-6}$ sec, where another detector counted 412 μ /hour. Thus 151 muons decayed in flight.

Now, applying the formula above, the number of muons/hour that we would expect to detect at sea level is much lower, namely

$$N(t) = 563e^{-(6.4 \times 10^{-6} s)/(2.2 \times 10^{-6} s)} = 31\mu / hr .$$

The conclusion that Frisch & Smith (1963, p. 348) draw from their experiment is that “the mesons decay much more slowly when they are in rapid flight, relative to us, than they do when they are at rest, with respect to us.”

The question is: How can this important difference between the calculated 31 μ /hour and the observed 412 μ /hour be explained?

The answer to this puzzling situation can be found within the framework of special relativity. Indeed, if we call respectively O and O' two observers that detect the same phenomenon, of which O is fixed on the ground and O' moves jointly with the event and measures a time T' , corresponding in our case to τ , also called *proper time*, the corresponding Lorentz transformation equation of special relativity $T = \gamma T'$, called *time dilation* (Rivadulla 2003, pp. 140-141) – where $\gamma = \frac{1}{\sqrt{1 - (v/c)^2}}$ –, determines that

$$T = \frac{2.2 \times 10^{-6} s}{\sqrt{1 - (0.9952)^2}} = 2.25 \times 10^{-5} \text{ sec},$$

that is

$$N(t) = 563e^{-(6.4 \times 10^{-6} s)/(2.25 \times 10^{-5} s)} = 424\mu / hr .$$

This prediction agrees quite well with the number of muons/hour detected experimentally. *The relativistic phenomenon of time dilation offers a satisfactory theoretical explanation of the observations.* In any case, time dilation – and its complement, length contraction – are theoretical novelties of first magnitude anticipated by special relativity. This is an excellent example of how a theoretical innovation makes the scientific explanation of an experimental result possible.

3 Innovation and inter-theoretical explanation

As I already stated in Rivadulla (2018, p. 380) “Productive reasoning serves both for *innovation* and *explanation* in the methodology of physics”. In what follows I will take the origin of wave mechanics as a case study that exemplifies the role that productive reasoning plays in invention, discovery, innovation or scientific creativity, and therefore also in explanation, in the theoretical natural sciences.

Recalling his own line of thinking, De Broglie (1941, §12.1. My emphasis, A.R.) acknowledges that the hypothesis of the “quanta of light” was the first bridge that was drawn in 1905 between the wave and corpuscular images of light:

In this hypothesis it is admitted that to any monochromatic plane light-wave of frequency ν corpuscles must be associated, the photons, which advance in the direction of the light propagation and that have an energy E and a momentum p given by the relations:

$$E = h\nu \qquad p = \frac{h\nu}{c} \qquad (12.1)$$

(...)

Relationships (12.1), which establish a link between the wave image (characterized by the frequency and direction of propagation of a monochromatic wave) and the corpuscular image (characterized by the energy and the momentum of a particle in uniform motion), suggest that, at least in the case of photons, it is appropriate to construct a synthetic theory in which the images of waves and corpuscles manage to find their place, insofar as they are exact. *And, in general, we get the idea that the same could happen for all kinds of particles that atomic physics has taught us to know.”*

To anticipate the wave image of matter, Louis De Broglie (1925) combined results from three theories: optics, special relativity and quantum physics, thus making use of *productive reasoning*: A monochromatic wave $\psi(x,t) = A(\sin kx - \omega t)$ with angular frequency $\omega = \frac{2\pi}{t}$ and wave number $k = \frac{2\pi}{\lambda}$ propagates with *phase velocity* $u_\varphi = \frac{\lambda}{t} = \frac{\omega}{k}$, i.e. $u_\varphi = \lambda\nu$. But when monochromatic waves of different wavelengths travelling in the same direction interfere, the resulting wave propagates with *group velocity* $v_g = \frac{d\omega}{dk}$.

Let us now suppose a material particle moving away with velocity $v=v_g$. Its total energy is, according to special relativity theory, $E = \gamma mc^2$, where $\gamma = (1 - \beta^2)^{-1/2}$, taking $\beta = v/c$. Let us now suppose that Einstein’s wave-particle duality of radiation is extensible to material particles. Then, according to quantum physics, the energy of the particle should be $E = h\nu$. Equating both values for E , we get $\nu = \frac{\gamma mc^2}{h}$. And applying the definition of phase velocity, we obtain $\lambda = u_\varphi t = \frac{u_\varphi h}{\nu} = \frac{u_\varphi h}{\gamma mc^2}$.

Now, as the phase and group velocities verify $u_\varphi = \frac{c}{\beta}$, and that $v_g = \beta c$ (De Broglie 1925, pp. 38-39 and 113), it results $u_\varphi v_g = c^2$. Substituting in $\lambda = \frac{u_\varphi h}{\gamma m c^2}$ formula $u_\varphi = \frac{c}{\beta}$ –and keeping in mind that $p = \gamma m v$ is the momentum of the material particle– we get $\lambda = \frac{h}{p}$, which is the well-known expression of the wave image of matter.

Only two years after, Clinton Davisson and Lester Germer confirmed experimentally De Broglie's theoretical innovation, which provided the theoretical explanation of the electron diffraction:

Mr. Davisson and Mr. Germer,..., were honoured to have discovered in 1927 the existence of *the diffraction of electrons by crystals*... Thus was established the existence of this beautiful phenomenon whose simple announcement, a few years before, would have provoked the incredulity of the physicists. And so was provided a magnificent direct confirmation of *the existence of a double corpuscular and undulatory aspect of the electron*. (De Broglie 1941, Chapter 2. My emphasis, A.R.)

4 Conclusion

Theoretical physics is entrusted with two tasks that can not be waived: to explain known results and to anticipate novel results. As the examples presented above show, theoretical innovation is explanatory. De Broglie's revolutionary proposal gives a theoretical explanation of the diffraction of material particles; and the explanation of time dilation is unthinkable without the incorporation of first rank theoretical novelties. Indeed, explanation and innovation are very closely related to each other.

ACKNOWLEDGMENTS. Con el recuerdo entrañable de Virginia, mi amada esposa.

I am very grateful to doctors Pablo and Diego Soler for careful reading of, and helpful comments on, a previous version of this paper.

References

- [1] De Broglie, L. (1925). *Recherches sur la Théorie des Quanta*. Annales de Physique 10^e Série, Tome III. Reprinted in Annales de la Fondation Louis de Broglie, Vol. 17, N^o 1, 1992.

- [2] De Broglie, L. (1941). *Continu et Discontinu en Physique Moderne*, Albin Michel, Paris.
- [3] Frisch, D. H. & Smith, J. H. (1963). "Measurement of the Relativistic Time Dilation Using m-Mesons", *American Journal of Physics*, 31, pp. 342-355.
- [4] Hempel, C. G. (1965). *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. New York: The Free Press.
- [5] Ostlie, D. A. and Carroll, B. W. (1996). *An Introduction to Modern Stellar Astrophysics*. Reading, MA: Addison-Wesley Publ. Co. Inc.
- [6] Rivadulla, A. (2003). *Revoluciones en Física*. Madrid: Editorial Trotta.
- [7] Rivadulla, A. (2004). *Éxito, Razón y Cambio en Física. Un enfoque instrumental en teoría de la ciencia*. Madrid: Editorial Trotta.
- [8] Rivadulla, A. (2005). "Theoretical Explanations in Mathematical Physics", in G. Boniolo *et al.* (eds.), *The Role of Mathematics in Physical Sciences*, pp. 161-178, Dordrecht: Springer,.
- [9] Rivadulla, A. (2008). "Discovery Practices in Natural Sciences: From Analogy to Production", *Revista de Filosofía*, vol. 33, nº. 1, pp. 117-137.
- [10] Rivadulla, A. (2010a). "Two Dogmas of Structural Realism. A Confirmation of a Philosophical Death Foretold", *Crítica. Revista Hispanoamericana de Filosofía*, vol. 42, nº. 124, pp. 3-29.
- [11] Rivadulla, A. (2010b). "Complementary Strategies in Scientific Discovery: Abduction and Production", in M. Bergman *et al.* (eds.), *Ideas in Action: Proceedings of the Applying Peirce Conference*, pp. 264-276, Nordic Studies in Pragmatism 1, Nordic Pragmatism Network, Helsinki.
- [12] Rivadulla, A. (2018). "Theoretical innovation. A new look at creativity in the natural sciences", in Cristian Saborido, Sergi Oms & Javier González de Prado (eds.): *Proceedings of the IX Conference of the Spanish Society of Logic, Methodology and Philosophy of Science*. Madrid: UNED.
- [13] Rivadulla, A. (2019). "Causal Explanations: Are They Possible in Physics?" In M. R. Matthews (ed.), *Mario Bunge: A Centenary Festschrift*, pp. 303-328. Cham: Springer.

Minas a cielo abierto en la comarca de Olivenza: las controversias científico-tecnológicas públicas

Álvaro Rodríguez*
University of Salamanca

Esta comunicación pretende sondear el debate sobre la minería a cielo abierto en la comarca de Olivenza, Extremadura (Proyecto Minero-Industrial Alconchel) desde el panorama teórico de las controversias científico-tecnológicas públicas. Esta clase de controversias se caracterizan bajo cuatro elementos: la existencia de un producto, teoría o práctica científico-tecnológica en torno a la cual se produce el debate; la presencia de expertos situados a ambos lados de la contienda; la apertura de diferentes interpretaciones conflictivas de los datos que se sopesan; y la transcendencia hacia el público del conflicto originado. Todo ello se encuentra aliñado por desacuerdos políticos y morales que condicionan las posturas. También intervienen una amplia cantidad de actores científicos y no científicos: la deriva de estas disputas hacia lo social implica factores que se encuentran fuera de las revistas, discusiones y laboratorios estrictamente científicos (Martin, 2014). El caso que nos ocupa se basa en la corta minera: un tipo de explotación minera con forma de cantera en cono invertido escalonado desarrollada a través de grandes excavaciones con maquinarias y explosivos.

El proyecto minero comienza en 2004 cuando la Dirección General de Ordenación Industrial, Energía y Minas de la Junta de Extremadura acepta el permiso de investigación llamado “Gato Montés” que más tarde recogería Valoriza Minería, una filial de la empresa española Sacyr, con el objetivo de obtener cobre, hierro y oro en el Dominio Geológico Norte de la Antiforma Olivenza-Monesterio, dentro de la zona llamada “Ossa Morena” que se extiende desde Portugal hasta el Guadalquivir. En 2015, la empresa propone una Concesión de Explotación Derivada a la Junta de Extremadura. “Gato Montés” (SIGEO, 2017) se incluye en la Ley 16/2015 de 23 de abril, de protección ambiental de la Comunidad Autónoma de Extremadura y quedó anunciada en el DOE número 248 de 29 de diciembre de 2017.

Podemos mencionar tres agentes en este caso: 1) la Junta de Extremadura y órganos regionales como la Administración Minera de Extremadura, el Sistema de Información Geológico Minero de Extremadura (SIGEO), o la Dirección General de Sostenibilidad y Medio Ambiente; 2) la empresa española Sacyr que defiende y orquesta la construcción de la mina; y 3) la plataforma de ciudadanos afectados “Alconchel y comarca de Olivenza

* alvarorr@usal.es

sin minas”. Estos tres agentes sustentan sus alegaciones en base a evidencias y marcos legislativos, además de ser asesorados por diferentes expertos científicos como geólogos, ambientalistas, ingenieros, economistas o biólogos. Otros actores relevantes son los partidos políticos que condicionan las interpretaciones populares y las opciones de resolución de estos debates.

Existe una particularidad: la especial omisión de información por parte de la empresa minera Sacyr y de la administración general; una situación, como menciona Martín (2014, p. 46) de *hidden research*. Esto dificulta enormemente el análisis: no existe un acceso público actual a los posibles estudios de impacto ambiental y social. De hecho, tampoco se ha llegado a expresar la autorización ambiental previa a la investigación y exploración minera por parte de la empresa, habiendo realizado sondeos diamantinos (DDH) en casi 16.000 metros hasta 2019 (Solaz Alpera, 2019). Así, el cuerpo de evidencias utilizado en este análisis se genera alrededor del informe “Análisis y Gestión de aguas”, que tampoco está disponible actualmente para el público.

El pasado febrero de este año el Ayuntamiento de Olivenza hizo saber por canales no oficiales que la Dirección General de Sostenibilidad de la Consejería para la Transición Ecológica y Sostenibilidad de la Junta de Extremadura había archivado la concesión de explotación de 2015 por la falta de documentación aportada por la empresa (HOY, 2021). Aun así, esto no estuvo acompañada de una Declaración de Impacto Ambiental negativa, o de una negación expresa del permiso de explotación, algo que deja abierta la posibilidad de futuros recursos por lo que podría ser prematuro dar por finalizada la controversia por completo.

Esta controversia se desarrolla en diferentes planos: en el plano ambiental se centra en el riesgo a corto y largo plazo sobre el impacto en el entorno, ecosistema, fauna y contaminación; en el plano político está inmersa en disputas sobre objetivos políticos convenientes; en lo económico se concentra en el progreso, la inversión, el desarrollo productivo y los puestos de trabajo; respecto a lo social tiene que ver con la percepción pública y los efectos en la vida local de la ciudadanía, abordada a través de la demanda de participación y transparencia; y en el plano científico-tecnológico se centra en la interpretación de los datos técnicos.

En el caso actual, se pueden mencionar tres marcos formados por estas orientaciones: el técnico-económico, en el que se centran la empresa y el gobierno y que hace énfasis en la necesidad del desarrollo industrial, la creación de mercado internacional y el aumento de puestos de trabajo, defendiendo el gran potencial minero-industrial del área metalogénica de Extremadura; el científico-tecnológico, utilizado para la justificación o crítica de los riesgos; y el ambiental-social, desde donde se sitúan los ambientalistas y la plataforma ciudadana recalando, tanto el impacto ecológico, como el social (expropiaciones obligadas o contaminación sonora).

Los argumentos técnicos en este conflicto se encuentran sumidos en la ambigüedad por la falta de documentación pública oficial. Acudiendo a los fragmentos disponibles del Análisis y Gestión de Aguas y la información recopilada por los movimientos locales y por los órganos regionales, los datos se pueden agrupar en tres conjuntos: contaminación ambiental, daños para la salud y explotación de acuíferos subterráneos.

Los datos sobre el daño a la salud se encuentran enlazados con la contaminación ambiental y son defendidos por ambientalistas y plataforma. Se centran en el fuerte impacto sobre el ecosistema del medio, daños en la superficie terrestre, flora y fauna, contaminación del aire por la generación de polvos tóxicos, de las aguas superficiales y de los acuíferos subterráneos por el drenaje ácido de residuos (Ecologistas en acción, 2019). Frente a estos riesgos, por parte de la empresa y el gobierno se hace referencia al seguimiento de los criterios de los técnicos, pero no se facilita información concreta sobre ello, algo que hace difícil el análisis. Para su argumentación, la empresa acude a las recientes orientaciones sostenibles recogidas en el Dictamen del Comité Económico y Social Europeo sobre “La minería no energética en Europa” (2009/C, 27/19), abogando por el control constante de los niveles de contaminación y la restauración de las zonas mineras una vez hayan cumplido su función.

El cuerpo de evidencias más consistente es el referido a la explotación de acuíferos subterráneos: es aquí donde encontramos el desacuerdo más explícito entre expertos. Según el Análisis sobre la Gestión de Agua, para la mina se accedería al acuífero cárstico de rocas calizas “Zafra-Olivenza” cuya condición sobrepasaría presumiblemente las necesidades del Proyecto sin consecuencias problemáticas, según afirman asesores de “FRASA Ingenieros Consultores” asociados a la empresa. Sin embargo y a partir de los datos de otras minas similares, desde la plataforma se resalta la sobreexplotación del acuífero y la repercusión en zonas protegidas por la llamada Red Natura 2000. Por parte de técnicos asesores en hidrología del Instituto Geológico y Minero de España (IGNE), la explotación continua de este acuífero de aguas fósiles frente a tres años de pocas lluvias produciría una bajada del nivel freático, la sequía de muchos pozos tradicionales y serios problemas en el abastecimiento de poblaciones como Olivenza, Alconchel, Táliga e Higuera de Vargas.

La existencia de estas perspectivas hace imprescindible la licencia política y social imponiendo un criterio mínimo: el de transparencia y fidelidad en la comunicación de información. Aquí no solo interfieren cuestiones científico-tecnológicas, también sociales relacionadas con los beneficios, los riesgos, la ética y la política, haciendo que la apelación a los expertos como mecanismo de resolución único sea insuficiente. Finalmente, también se puede encontrar un elemento añadido que dificulta especialmente el campo epistémico: la incertidumbre. Este es un problema que recae en la rigidez de las predicciones y argumentos científicos y en la de las soluciones y medidas políticas. Debido a ello

y frente a la infradeterminación de los datos por la teoría consecuente se hace imprescindible la atención a la diversidad de perspectivas y a la intervención de posibles valores e intereses que condicionen las interpretaciones.

Referencias

- [1] Ecologistas en Acción (2019). Informe: Minería Especulativa en España. Versión 1.0. Madrid.
- [2] Eldiario.es (2020). “Fiebre del cobre en el suroeste de Badajoz, la polémica entre proyectos sostenibles o perjudiciales”, 26 de enero.
- [3] Hoy (2021). “La Junta rechaza el proyecto de mina de Alconchel”. 2 de febrero.
- [4] Martin B. (2014). Controversy Manual. Irene Publishing. Creative Commons. ISBN 978-1-291-67241-1.
- [5] Sigeo (2017). Información Pública Aprovechamiento Recurso Sección C) “Gato Montés”, Expte. n.º 06C12606-10.
- [6] Solaz Alpera, I. (2019). Alconchel: un nuevo yacimiento de CU-AU en el SO de la Península Ibérica. Valoriza Minería.

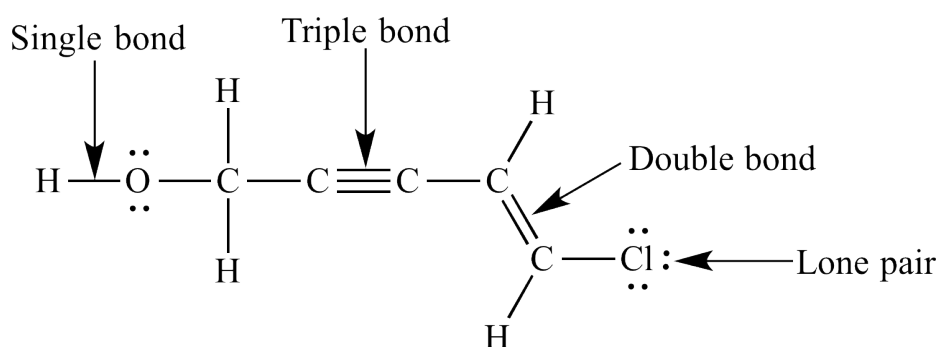
An inferentialist account of Lewis structures

Pedro J. Sánchez Gómez*
Universidad Complutense de Madrid

Mauricio Suárez**
Universidad Complutense de Madrid

Introduction

Lewis structures (LS) conspicuous examples of scientific representations. They are schematic descriptions of the structure of molecules in terms of the valence electrons of the atoms that compose them. A typical LS looks like this:



To understand the role of LS, a distinction must be drawn between the part of chemistry that tries to address the question “What makes a chemical statement true?” and that part that simply aims at providing insightful accounts of chemical phenomena. The first program is characteristic of what is usually called *physical chemistry* (PhysChem), a wide field that includes quantum chemistry, statistical chemical thermodynamics and other sub-fields such as chemical spectroscopy or electrochemistry. The second program is typical of what Joachim Schummer has called, in happy expression, “*the chemical core of chemistry*” (Schummer, 1998). We shall refer to this specifically chemical approach as *chemical chemistry* (ChemChem). The role of LS’s in PhysChem is ancillary: They are employed as pictograms to represent the different chemical species, but they play no role in the theorizing within PhysChem. On the other hand, ChemChem is the realm of LS, where they are instrumental to any explanation.

In this work we shall focus on ChemChem. We shall study the conditions that enable the explanatory power of LS’s within an inferentialist view of chemical reactivity.

* pedros@edu.ucm.es

** msuarez@filos.ucm.es

Chemical explanation: reaction mechanisms

The explanation of a reaction within ChemChem is closely linked to the notion of *reaction mechanism*. A reaction mechanism is “a specification, by means of a sequence of elementary chemical steps, of the detailed process by which a chemical change occurs” (Lowry and Richardson, 1987, p. 190. Cited in Goodwin, 2017). The usual way of presenting a mechanism is graphical, in terms of LS’s. Let us see it in a standard organic chemistry example, the reactions of electrophilic aromatic substitution (**Fig. 1**).

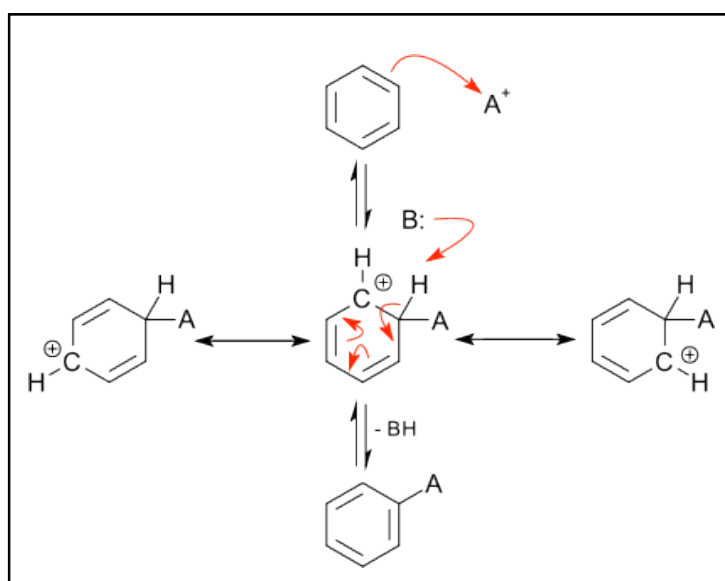


Figure 1: Electrophilic Aromatic Substitution.

License notice for the image: V8rik at English Wikipedia

(<https://commons.wikimedia.org/wiki/File:BasicAromaticSubstitution.svg>),

“BasicAromaticSubstitution”, <https://creativecommons.org/licenses/by-sa/3.0/legalcode>

This form of graphical explanation also has some predictive value. For example, it can be used to study the influence of a substituent group on the relative yields when more than a product is generated. For example, an *amino* group ($-NH_2$) directs the reaction to the *ortho* and *para* products, by means of a resonant stabilization of the cationic aromatic intermediate (**Fig. 2**). On the other hand, in the reaction on the meta position the positive charge cannot be delocalized over the amino group, thus being less stable (**Fig. 3**). As a conclusion, the yields for the ortho and para substitutions are expected to be similar and substantially greater than that for the meta product. Such a prediction can be extended to any *activating* groups and has been being experimentally confirmed for more than a century.

Similar arguments can be found in any introductory textbook for any organic reaction. Moreover, a quick look on the (countless) current chemical research publications shows that this explanatory scheme is far from being a mere introductory aid with didactic purposes.

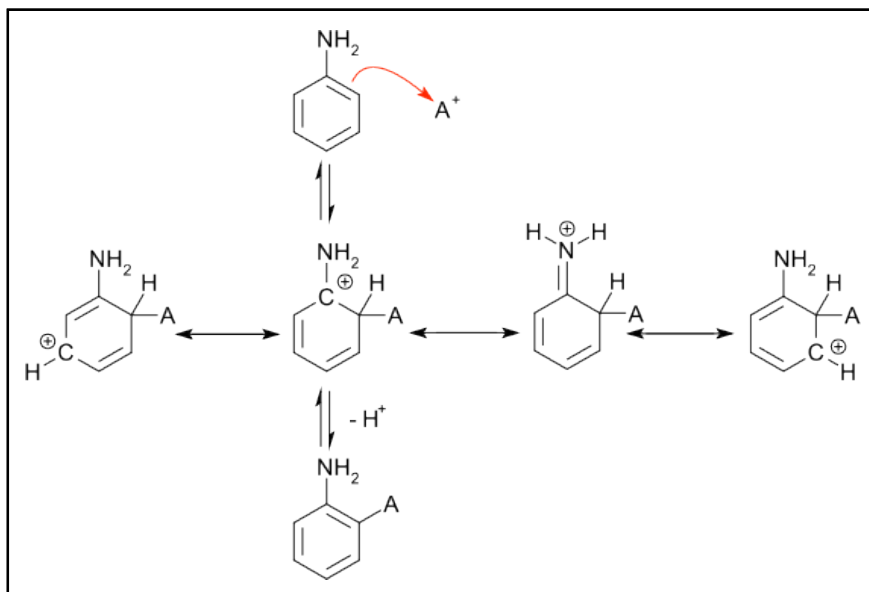


Figure 2: Ortho electrophilic substitution on aniline.

License notice for the image: V8rik at English Wikipedia
 (<https://commons.wikimedia.org/wiki/File:ElectrophilicAromaticSubstitutionOrthoDirectors.svg>),
 “ElectrophilicAromaticSubstitutionOrthoDirectors”,
<https://creativecommons.org/licenses/by-sa/3.0/legalcode>

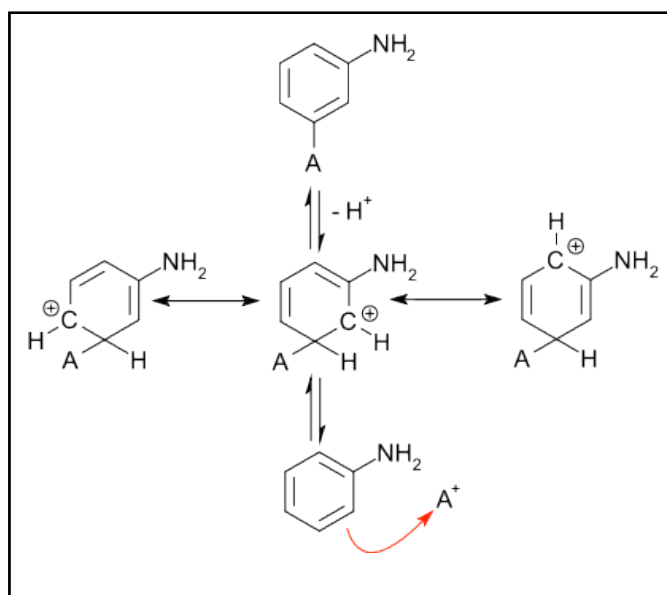


Figure 3: Meta electrophilic substitution on aniline.

License notice for the image: V8rik at English Wikipedia
 (https://commons.wikimedia.org/wiki/File:EAS_meta_deactivation.svg),
 “EAS meta deactivation”, <https://creativecommons.org/licenses/by-sa/3.0/legalcode>

Lewis' structures as stereotypes

LS's, together with a rather restricted set of rules to *operate on* them, allows us to infer the reactivity of a substance. The explanatory value of these inferences is evident. Observable features, such as the kinetics of a reaction, or its stereospecificity, can be explained in terms of the structural features in the LS's of the intermediate species of the mechanism. On the other hand, the predictive power of the inferences based on LS's is limited. At best, and only in some model reactions, some qualitative conclusions can be achieved about the relative order of the yields of competing reactions. For example, the electrophilic aromatic substitution in aniline that we have reviewed in the previous section. On the other hand, such clear predictions can seldom be produced for most realistic chemical reactions (Jenck 1981).

In sum, the role of LS's within chemical reasoning is somehow contradictory. On the one hand, they are instrumental to the way chemists explain reactions. On the other, they provide a limited degree of control over most of these reactions. Our claim is that this tension can be accounted for within an inferentialist approach to LS. In this view, for a scientific representation to be effective, some intended uses of an agent are required, and this implies some pragmatic considerations in the interpretation of the model sources (here the LS themselves). There are also pragmatic skills involved in the appropriate descriptions of the targets and the means by which sources can be employed to draw inferences regarding the target (Suárez, 2004). Yet, literally instantiating the structure of a molecule, as it occurs in nature, as it were regardless of description, is not within the purview of scientific representations like LS's.

A possible account of the semantics of LS's is thus as *stereotypes*, in Putnam's (1975) influential account. According to Putnam, a stereotype about an object X is "a conventional (frequently malicious) idea (which may be wildly inaccurate) of what X looks like or acts like or is" (Putnam 1975, p. 249). Stereotypes guarantee a semantic competence only in some standard contexts, where speakers have been appropriately trained. Otherwise, they are usually idiosyncratic, often misguided, and thus they do not qualify as the meaning of a word. In this view, the speakers implicitly defer to their respective communities of experts to determine the meaning of a natural kind term.

LS's are highly idealized models of the structure of molecules, yet they enable *chemical chemists* to draw, and share, inferences about reactivity. This pragmatic side is linked to a semantic deference towards the community of physical chemists. Chemical chemists, sometimes even explicitly, defer the task of fixing the actual properties of substances to quantum chemists, spectroscopists and the so. It is in fact common in chemistry texts to quantify structural features derived from LS's, such as bond lengths or angles, with data deduced from spectra or from quantum chemistry calculations (there are countless examples of this approach such as the tables included in Linares, Braïda and Humbel, 2006).

Without this deference, ChemChem would collapse into an ad hoc collection of models of reactions. However useful as representations, such models would be lacking in predictive power.

Conclusion

We have advanced an inferentialist account of Lewis structures (LS's). In this framework, the tension between the sound explicative value of LS's and their poor predictive power can be readily accounted for. Besides, the notion of semantic deference, linked to that of stereotypes, provides a *natural* interpretation of the relation between *Chemical Chemistry*, the part of chemistry explicitly based on LS's, and *Physical Chemistry*, that is, the application of physical theories to chemical phenomena.

References

- [1] Goodwin, W. (2017). The origins of the reaction mechanism. In Glennan, S. and Illari, P. (Eds.), *The Routledge Handbook of Mechanisms and Mechanical Philosophy (1st ed.)*. Routledge. <https://doi.org/10.4324/9781315731544>
- [2] Jenck, W. P. (1981). How Does a Reaction Chooses its Mechanism? *Chemical Society Review*, 10, 345-375.
- [3] Linares, M., Braïda, B and Humbel, S. (2006). Lewis-Based Valence Bond Scheme: Application to the Allyl Cation. *Journal of Physical Chemistry A*, 110, 2505-2509.
- [4] Lowry, T.H. and Richardson, K.S. (1987). *Mechanism and Theory in Organic Chemistry* (3rd ed.). Harper & Row: New York.
- [5] Putnam, H. (1975). The Meaning of 'Meaning'. *Minnesota Studies in the Philosophy of Science*, 7, 131-193. Repr. in *Philosophical Papers Vol. 2: Mind, Language and Reality* (1975), 215-271. Cambridge: Cambridge University Press.
- [6] Schummer, J. (1998). The conceptual core of chemistry. A conceptual approach I. *HYLE -International Journal for Philosophy of Chemistry*, 4, 129-162.
- [7] Suárez, M. (2004). An Inferential Conception of Scientific Representation. *Philosophy of Science*, 71, 767-779.

Pluralism beyond complexity: Reconceptualizing the disunity of Psychiatry

Mariano Sanjuán*
Autonomous University of Madrid

Among all the questions that have long concerned philosophers of science, one stands as particularly important: are the sciences one or many? Philosophers of science have traditionally developed two straight answers. On the one hand, some have championed that unification is the main aim of science as the history of science is jump-packed with cases of theoretical convergence. They declare that although currently many areas of science cannot be conciliated, it is a matter of time that they converge into one unique account. Among others, the two most relevant arguments for it have been some version of epistemic reductionism and metaphysical fundamentalism. Epistemic reductionists claim that either or both the laws and concepts of the special sciences can be derived from one more fundamental theory, and it is an aim of science to find it (Nagel 1961; Klein 2009), while fundamentalists claim that scientific disciplines follow a unidirectional and hierarchically structured organization divided by levels of fundamentality, with particle physics located at its base, where more fundamental levels of aggregation are fully responsible for the happenings in higher levels (Oppenheim & Putnam 1958). Despite criticism, unification is still a considerably defended approach to the development of the sciences and a popular stance towards physics and highly mathematized theories (Tahko 2021).

On the other hand, some believe that cases of disunity are at least as frequent as those of unification, and therefore it is likely that the sciences remain as disunified as they look now. This view is commonly dubbed scientific pluralism (Keller et al. 2006), a normative philosophical standpoint that supports the cultivation and proliferation of multiple systems of practices and beliefs (Chang 2012). As it emerged as a reaction against unificatory tendencies, pluralism has been defended on the basis of anti-reductionism and anti-fundamentalism. Anti-reductionism implies that either or both the laws and concepts of different theories cannot be fully translated nor reduced to any alternative theory, while anti-fundamentalism holds that it is a mistake to regard lower levels of aggregation as fully and unidirectionally responsible of the dynamical happenings in higher levels. As a result, anti-fundamentalists claim that nature is dappled (Cartwright 1999) and disordered (Dupré 1993), so must be our knowledge of it. Pluralists vindicate the autonomy and distinctiveness of methods, languages and theories across disciplines, a view that is

* marianosanjuansalinas@gmail.com

especially widespread among philosophers of biology and medicine. As a medical field, Psychiatry is commonly regarded as a pluralistic field (De Vreese 2010; Kendler 2012; Tekin & Machery 2019). Psychiatry, it is said, provides an exemplary case of non-reductionist and non-foundationalist science.

Two main areas within Psychiatry are deemed pluralistic: explanation (Kendler 2005; van Bouwel 2014; Gijsbers 2016) and classification (Tsou 2015; Bueter 2019). Explanatory pluralism claims that coexisting explanations at multiple levels of aggregation are required to shed light on the causal story of mental diseases. As for taxonomical pluralism, it privileges coexisting systems of classification (taxonomies) as a better understanding of mental diseases. Both explanatory and taxonomical pluralism draw on the idea that mental diseases are the product of complex systems (Mitchell 2009). On this occasion, I inquire whether complexity is sound ground enough to build pluralism upon. Note that I do not question whether mental diseases are truly complex; I concede they are. My concern is subtly different: assuming that pluralism should be vindicated, is appealing to the complex nature of mental diseases the most advantageous argumentative strategy? In what follows, I will argue that pluralist perspectives should not fall back on considerations on the putative complex nature of mental diseases. I hold that pluralism about Psychiatry constitutes an encouraging perspective, but it calls for a more vigorous conceptual justification since accounting for scientific pluralism by appealing to complexity is not compelling enough.

My talk is divided into four parts. Firstly, I address some methodological concerns regarding the biomedical and the biopsychosocial models in Psychiatry. I conclude that both face difficulties that show that even the very foundational issues of Psychiatry should be managed pluralistically. In the following section, I explore explanatory and taxonomical pluralism through a case study –the correlation between depressive disorders and 5-HTTLPR polymorphism, addressed in Mitchell (2009)– and the critiques to the DSM-5, respectively. I conclude that both taxonomical and explanatory pluralists agree on considering mental diseases as complex phenomena or products of complex systems. As a consequence, they endorse pluralism by means of the alleged complex character of mental diseases. After that, in the third part, I raise three objections to those who endorse pluralism solely relying upon the concept of complexity: (i) complexity circumvents accurate definition, (ii) whether a system is complex depends on our epistemic contexts (Ruphy 2016) and (iii) pluralists' conception of complexity is unnecessarily ontologically committed. Regarding this latter point, I emphasize that, by directly appealing to nature's properties, pluralists bring about the same inferential move they reject as mistaken in reductionist accounts: concluding philosophical standpoints from metaphysically non-justified ideals about nature's properties. As an alternative, I propose a metaphysically deflationary and

pragmatic account of the disunity of science to ground pluralism. For that, I rely on Otto Neurath's idea of "cooperative science".

In the last section, I present Neurath's take on the disunity of science, basing on Potochnik (2011) and Neurath himself. Although apparently contradictory, since Neurath is commonly taken as the prominent impeller of the unity of science movement, some studies have shown that there are many conceivable "unities of science", and Neurath's account is indeed closer to pluralism than usually regarded (Cartwright et al. 1991; Symons, Pombo & Torres 2011; Cat & Tuboly 2019). Neurath's views on the Unity of Science aim to pursue the practical and harmonic integration between systems of knowledge rather than reducing their respective "jargons" into one single account (Hacking 1996). Thus, a Neurathian account of the disunity of science should take on how different disciplinary perspectives cooperate to address specific scientific problems independently of unificatory desiderata. I hold that this is all the more emphatic in the case of mental diseases. Reconceptualizing the disunity of Psychiatry in that vein could preserve the vindication of pluralism in Psychiatry, saving all the conceptual difficulties that the idea of complexity entails. I will support an empirically oriented, pragmatic and metaphysically-deflationary version of pluralism stemming from Neurath's work, showing how Psychiatry provides an insightful example of intertheoretical scientific cooperation, and how pluralism can be endorsed beyond the alleged complex nature of mental diseases.

References

- [1] Bueter, Anke (2019). A Multi-Dimensional Pluralist Response to the DSM-Controversies. *Perspectives on Science* 27 (2), 316-343.
- [2] Cartwright, Nancy (1999). *The Dappled World*. Cambridge: Cambridge University Press.
- [3] Cartwright, Nancy et. al. (1991). *Otto Neurath: Philosophy Between Science and Politics*. Cambridge: Cambridge University Press.
- [4] Cat, Jordi & Tuboly, Adams T. (2019) (eds.). *Neurath Reconsidered*. Dordrecht: Springer. Chang, Hasok (2012). *Is Water H₂O?* Dordrecht: Springer.
- [5] De Vreesse, Leen, Weber, Eric & Van Bouwel, Jeroen. (2010) Explanatory pluralism in the medical sciences: Theory and practice. *Theor. Med. Bioeth.* 3, 371-390.
- [6] Dupré, John (1993). *The Disorder of Things*. Harvard: Harvard University Press.
- [7] Gijsbers, Victor (2016). Explanatory Pluralism and the (Dis)Unity of Science: The Argument from Incompatible Counterfactual Consequences. *Frontiers in Psychiatry*. 7 (32), 1-10.

- [8] Hacking, Ian (1996). The Disunities of the Sciences, in Peter Galison & David J. Stump (eds.). *The Disunity of Science*. Stanford: Stanford University Press, 37-74.
- [9] Kellert, Stephen H., Longino, Helen E. & Waters, C. Kenneth (eds.) (2006). *Scientific Pluralism*. Minnesota: The University of Minnesota Press.
- [10] Kendler, Kenneth S. (2005). Toward a Philosophical Structure for Psychiatry. *American Journal of Psychiatry* 162, 433-440.
- [11] Kendler, Kenneth S. (2012). The dappled nature of causes of psychiatric illness: replacing the organic- functional/hardware-software dichotomy with empirically based pluralism. *Molecular Psychiatry* 17, 377-388.
- [12] Klein, Colin (2009). Reduction without Reductionism: A Defence of Nagel on Connectability. *The Philosophical Quarterly* 59 (234), 39-53.
- [13] Mitchell, Sandra D. (2009). *Unsimple Truths*. Chicago: The University of Chicago Press.
- [14] Nagel, Ernst (1961). *The Structure of Science: Problems in the Logic of Scientific Explanation*. New York: Harcourt.
- [15] Oppenheim, Paul & Putnam, Hilary (1958). Unity of Science as a Working Hypothesis. *Minnesota Studies in the Philosophy of Science* 2: 3-36.
- [16] Potochnik, Angela (2011). A Neurathian Conception of the Unity of Science. *Erkenntnis* 74 (3), 305- 319.
- [17] Ruphy, Stephanie. (2016). *Scientific Pluralism Reconsidered*. Pittsburgh: Pittsburgh University Press. Symons, John, Pombo, Olga & Torres, Juan Manuel (eds.) (2011). *Otto Neurath and the Unity of Science*. Dordrecht: Springer.
- [18] Tahko, T. (2021). *Unity of Science*. Cambridge University Press.
- [19] Tekin, Serif & Machery, Eduoard (2019) (eds.). Psychiatry and Its Philosophy. Special Issue in *Synthese*, 196.
- [20] Tsou, Jonathan Y. (2015). DSM-5 and Psychiatry's Second Revolution: Descriptive vs. Theoretical Approaches to Psychiatric Classification, in Steeves Demazeux and Patrick Singy (eds.). *The DSM-5 in Perspective: Philosophical Reflections on the Psychiatric Babel*. Dordrecht: Springer.
- [21] Van Bouwel, Jeroen (2014). Pluralists About Pluralism? Different Versions of Explanatory Pluralism in Psychiatry, in Maria Carla Galavotti et al. (eds.). *New Directions in the Philosophy of Science*. Dordrecht: Springer, 105-119.

Contra la redundancia

Albert Solé*
Universitat de Barcelona

Algunos partidarios de la interpretación de Everett o de los muchos mundos [MM] han argumentado que la mecánica bohmiana [MB] incorpora una estructura de muchos mundos en la función de onda y que, en consecuencia, la adición de los corpúsculos a dicha ontología resulta superflua y redundante. Siguiendo a Callender (2010), me referiré a esta línea de ataque como el “argumento de la redundancia” en contra de MB. En este artículo, voy a revisar dicho argumento a la luz del reciente debate sobre la interpretación de la función de onda en MB, argumentando que su conclusión no está debidamente justificada y ofreciendo nuevas consideraciones en favor de MB.

Wallace (2003) ha articulado la que quizás sea la versión más sofisticado del argumento de la redundancia, haciendo uso de lo que el autor denomina como “Criterio de Dennett” [CD]:

A macro-object is a pattern, and the existence of a pattern as a real thing depends on the usefulness –in particular, the explanatory power and predictive reliability– of theories which admit that pattern in their ontology. (Wallace 2003, p. 93)

Para mencionar un ejemplo recurrente, CD asegura que cualquier patrón que funcione como un gato (de acuerdo con nuestra mejor ciencia sobre gatos) es, en realidad, un gato. Por supuesto, los partidarios de MM deben considerar que los gatos son patrones discernibles en la estructura de la función de onda, ya que asumen que la función de onda (interpretada como un campo físico) es lo único que existe a nivel fundamental. Suponiendo que el mecanismo de la decoherencia da lugar a una división apropiada de la función de onda en ramas, CD garantizaría que cada una de estas ramas constituye un mundo con un sinnúmero de objetos macroscópicos siempre y cuando pudiéramos encontrar ahí patrones con el perfil funcional adecuado. Ahora bien, estos mismos patrones forman parte de la ontología bohmiana al incluirse en la misma la función de onda. Así, CD implicaría que las ramas de la función de onda también constituyen muchos mundos en MB, sin importar cuál sea la configuración de los corpúsculos, los cuales podrían ser vistos como un engranaje ocioso que convendría eliminar mediante la afilada navaja de Ockham.

* Albert.Sole@ub.edu

En la literatura se han ofrecido varias respuestas al argumento de la redundancia. Entre ellas, hay intentos de mostrar que los patrones en la función de onda no exhiben el perfil funcional adecuado como para ser identificados con objetos macroscópicos (Lewis 2007) o que MM no proporciona una solución satisfactoria al problema de medida (Callender 2010). También se ha hecho notar que las interpretaciones *tridimensionalistas* de MB que no reifican la función de onda no pueden ser tildadas de redundantes. Estoy de acuerdo con esta última observación y la articularé a continuación. Sin embargo, en mi opinión, existe otra razón por la que incluso los bohmianos que reifican la función de onda interpretándola como un campo físico tampoco cometen redundancia. Dicha razón no ha sido debidamente advertida en la literatura y se basa en el hecho de que partidarios de MB y de MM dan descripciones muy diferentes de la función de onda universal, o campo cuántico, aun cuando coinciden en interpretarla como un campo físico. Esto permite colegir que la función de onda universal en MB (ψ_{MB}) se considera algo diferente a la función de onda universal en MM (ψ_E). Debido a esta diferencia en la interpretación, los defensores de MB pueden estar justificados en asumir que los gatos *no* son patrones en la función de onda bohmiana sin que esto implique que los defensores del MM estén equivocados al asumir que los gatos sí pueden ser patrones en la función de onda everettiana.

Consideraré en primer lugar aquellas interpretaciones de MB que no toman la función de onda como denotando un campo físico real. Es indiscutible que estas interpretaciones no pueden ser una diana del argumento de la redundancia, puesto si no hay campo cuántico en la realidad, no hay patrones en dicho campo y no cabe identificar objetos macroscópicos con tales patrones. Los defensores de la redundancia suelen descartar este tipo de interpretaciones tildándolas de instrumentalistas e, incluso, han argumentado que los bohmianos *deben* reificar la función de onda. Así, por ejemplo, Brown y Wallace (2005) consideran que esto es así dado que la función de onda es contingente, estructuralmente compleja y que está dotada de una evolución temporal no trivial.

Sin embargo, conviene notar que en la última década se han planteado nuevas interpretaciones tridimensionalistas de MB, diferentes de la interpretación nomológica de la función de onda a la que Brown y Wallace hacen referencia. Así, se ha argumentado que la función de onda representa propiedades disposicionales irreducibles de las partículas bohmianas o, alternativamente, fuerzas tridimensionales newtonianas (o aristotélicas). La moraleja que cabe extraer de este debate es que, para articular una interpretación de MB satisfactoria desde el punto de vista de la explicación, hay que asumir un cierto realismo respecto de la función de onda, pero ello *no* supone necesariamente interpretar la función de onda como un campo físico.

Consideraré, a continuación, las versiones de MB que sí interpretan a la función de onda como un campo físico. Lo primero que hay que tener en cuenta es que este campo tiene características bastante inusuales. Primero, está definido en un espacio N -dimensional

(N es el número de partículas del universo) conocido como “espacio de configuración.” Ninguna de las dimensiones de dicho espacio se corresponde con una dimensión del espacio tridimensional de nuestra experiencia. En segundo lugar, el campo actúa sobre las partículas de acuerdo con la llamada “ecuación de guía”, pero su propia evolución viene dada por la ecuación de Schrödinger y es independiente del movimiento de las partículas. En tercer lugar, su acción sobre las partículas depende de su forma, pero no de su amplitud. Dadas estas y otras peculiaridades del campo cuántico, Bohm y Hiley (1993) lo interpretan como un campo “no mecánico” y de “información.” A su vez, Valentini (1992) caracteriza la función de onda como un campo de “causa formal aristotélica.” Está claro que tal énfasis en el aspecto formal del campo se debe a que el aspecto material se reserva para los corpúsculos, los cuales se supone que transportan la masa, la carga, la energía, etc. Pero, dadas estas consideraciones sobre la naturaleza casi inmaterial del campo cuántico, no sorprende que los bohmianos rechacen considerar que éste pueda constituir objetos macroscópicos.

Conviene hacer dos consideraciones finales. En primer lugar, podría objetarse que la diferencia en la interpretación de ψ_{MB} y ψ_E está justificada puesto que, matemáticamente, ψ_{MB} y ψ_E son el mismo objeto. Estoy en desacuerdo. Es plausible suponer que la interpretación física de un término (teórico) depende de todas las leyes y axiomas fundamentales en los que figura dicho término y, por tanto, de las relaciones que guarda con otros términos. De hecho, esta sería una consecuencia de asumir una teoría holística del significado. En el caso que me ocupa, la ecuación de guía que relaciona el campo cuántico con la velocidad de los corpúsculos es un axioma fundamental de MB pero no figura en MM. Esta diferencia en la red conceptual de ambas teorías justifica la diferencia de interpretación y, de hecho, explica por qué los everettianos no proporcionan caracterizaciones del campo cuántico similares a las que acabo de mencionar.

El segundo lugar, cabe preguntarse si la estrategia defendida aquí equivale, o no, a una renuncia o violación de CD. ψ_E y ψ_{MB} tienen la misma estructura matemática. Si los defensores de MM insisten en que un gato debe de ser identificado con un patrón de amplitudes relativas del campo cuántico en el espacio de configuración, entonces hay que conceder que, si ψ_E instancia dicho patrón, ψ_{MB} también lo hace. Si, a pesar de ello, se asume los gatos no pueden estar constituidos por el campo bohmiano, esto equivale efectivamente a restringir CD. Pero se trata una forma *justificada* de hacerlo. Vale la pena mencionar, a este respecto, que incluso en Filosofía de la Mente, criterios funcionalistas como CD han sido a menudo restringidos de alguna manera para evitar resultados poco intuitivos. Al final, la cuestión en juego puede ser si procede un uso liberal y sin restricciones de CD o uno más chovinista (véase Callender 2010).

Referencias

- [1] Bohm, D; Hiley, B. (1993): *The Undivided Universe: An Ontological Interpretation of Quantum Theory*, London: Routledge & Kegan Paul.
- [2] Brown, H.; Wallace, D. (2005): “Solving the Measurement Problem: de Broglie-Bohm loses out to Everett”, *Foundations of Physics* 35: 517-540.
- [3] Callender, C. (2010): “The Redundancy argument in Bohm’s theory”. Manuscrito no publicado.
- [4] Lewis, P. (2007): “How Bohm’s Theory Solves the Measurement Problem”, *Philosophy of Science* 74, 749-760.
- [5] Valentini, A. (1992): *On the Pilot-Wave Theory of Classical, Quantum and Sub-Quantum Physics*. Tesis doctoral. International School for Advanced Studies, Trieste.
- [6] Wallace, D. (2003): “Everett and Structure”, *Studies in the History and Philosophy of Modern Physics* 34: 87-105.

Evolutionary Fitness: Propensities, Probabilities, and Experimental Statistics

Mauricio Suárez*
Universidad Complutense de Madrid

Evolutionary fitness is often regarded by many to be a probabilistic disposition, or propensity to reproduce successfully (see Brandon, 1978; Mills and Beatty, 1979). There is even a long standing propensity interpretation of fitness (PIF) which is part of a larger tradition in evolutionary thinking that takes fitness or adaptiveness to be a causally explanatory concept (Sober, 1984, 2001). Yet, there is little consensus as to the specific kind of propensity that fitness is. On the contrary, there are open questions in the field as to how to formally represent fitness, how exactly it is an explanatory concept, and what exactly it explains. Critics have been quick to latch onto such disagreements in order to argue that fitness is not causally explanatory after all (Walsh, Ariew, Mahen, 2016), that it does not reflect causal relations (Walsh, 2010), and that there are no propensities underlying adaptation phenomena in evolutionary biology (Ariew and Ernst, 2009).

The current impasse suggests that there are fundamental issues at stake regarding the nature of propensity and its explanatory power that are yet to be clarified. In this talk a more complex and nuanced framework is presented than is typically assumed for modelling chancy phenomena in general, the ‘complex nexus of chance’ (CNC). Contrary to what has been conventional in the philosophy of probability, this approach distinguishes clearly propensities from both probabilities and the finite frequency data that are used to test them (see e.g. Mellor, 1971; Suárez, 2011; 2015). It is argued that the application of this framework to fitness bears significantly on a few important problems currently discussed in relation with (PIF).

Firstly there is the problem concerning the exact formal or mathematical representation of fitness as propensity. The relevant discussion here broaches two technical aspects of statistical modelling, informing what are sometimes known as the moments problem and the delayed selection problem (Sober, 1984; Beatty and Finsen, 1989; Pence and Ramsey, 2013). On the one hand there is the demonstrable empirical fact that fitness is often sensitive to higher moments of the statistical distribution for reproductive success. Hence identifying fitness with just the statistical mean average (the expected value, or expectation) of a probability distribution will often miss out critical differences down the lineage. The differences can be so critical as to entirely reverse judgements of relative

* msuarez@filos.ucm.es

fitness between individual organisms (or traits, or genes – more about this later on). But the idea that fitness, understood as a propensity, must necessarily be identified with some or other moment of a probability distribution presupposes that all propensities are statistical functions, or formal moments of the distributions, which is nowadays questionable in the philosophy of probability, and it is indeed rejected by the CNC. I consequently suggest that CNC accounts for the statistical modelling of fitness without such assumptions, and thus delivers us from the problem of moments.

The second narrow technical issue concerns whether fitness is short or long term; i.e. whether it involves reproductive success in the most proximate generations, or perhaps even just the next generation; or whether, by contrast, fitness refers meaningfully only to reproductive success down the generations – or perhaps even hypothetical success in some infinite reproductive limit. On a propensity analysis, the issue may at first sight seem merely a version of the debate regarding ‘single case’ versus ‘long run’ propensity interpretations of probability. If so, the delayed selection problem would boil down merely to a difference regarding the appropriate type of propensity involved, where those advocating long term fitness would be implicitly if not explicitly adopting a ‘long run’ propensity account. However, I argue that these distinctions are in fact tangential. Long term fitnesses, in particular, are perfectly compatible with ‘single case’ propensities, as advocated by the CNC. This has consequences for the precise mathematical definitions that are appropriate when modelling fitness in different contexts, and whether or not they issue in contradictions.

Then there is the second and more general issue, namely the explanatory role of fitness. Advocates of the PIF typically defend the view that fitness is a causally explanatory property of biological entities – and for this reason they are sometimes known as ‘causalists’ (Abrams, 2012). Critics of the PIF by contrast, tend to view fitness as not particularly an explanatory concept – certainly not a causally explanatory one –, but rather a descriptive or generalising concept. In the last part of the talk, I argue that CNC shows both ‘causalists’ and ‘statisticalists’ to be in part right. Propensities are indeed explanatory entities, but in accordance to CNC they typically explain not frequencies in data, but the single case chances that they give rise to within particular chance set ups. Thus ‘fitness’ is indeed often a name for an explanatory propensity, but not merely that: it is also a name used for the distinct probability distribution within a statistical model that is adequate for the purpose of representing the single case chances manifested. And in turn these chances are used to account for the actual data recorded in observational studies of reproductive success, where ‘fitness’ is also sometimes confusingly used to refer to the finite frequencies in the data for reproductive success. The disambiguation of these three distinct but mutually related uses of fitness is essential for a better understanding of its explanatory power.

Selected References

- [1] Abrams, M. (2009), “The Unity of Fitness”, *Philosophy of Science*, 76, pp. 750-761.
- [2] Ariew, A. and R. C. Lewontin (2004), “The Confusion of Fitness”, *British Journal for the Philosophy of Science* 55, pp. 347-363.
- [3] Beatty, J. and S. Finsen (1989), “Rethinking the Propensity Interpretation: A Peek Inside Pandora’s Box”, in M. Ruse (ed.), *What the Philosophy of Biology is*, Kluwer Academic Publishers, pp. 17-30.
- [4] Brandon, R. N. (1978), “Adaptation and Evolutionary Theory”, *Studies in History and Philosophy of Science*, 9, pp. 181-206.
- [5] Brandon, R. N. (1990), *Adaptation and Environment*, Princeton University Press.
- [6] Crow, J. and Kimura, M. (1956), “Some Genetic Problems in Natural Populations”, *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, 4, pp. 1-22.
- [7] Drouet, I. and F. Merlin (2015), “The Propensity Interpretation of Fitness and the Propensity Interpretation of Probability”, *Erkenntnis*, pp. 457-468.
- [8] Gillespie, J. H. (1974), “Natural Selection for Within-Generation Variance in Offspring Number”, *Genetics*, 76, pp. 601-606.
- [9] Grimmett, G. and D. Stirzaker (1982), *Probability and Random Processes*, Oxford: Oxford University Press.
- [10] Humphreys, P. (1985), “Why Propensities Cannot be Probabilities”, *The Philosophical Review*, 94 (4), pp. 557-570.
- [11] Krzanowski, W. J. (1998), *An Introduction to Statistical Modelling*, London: Wiley and Sons.
- [12] Mellor, H. (1971), *The Matter of Chance*, Cambridge: Cambridge University Press.
- [13] Mills, S. and J. Beatty (1979), “The Propensity Interpretation of Fitness”, *Philosophy of Science*, 46 (2), pp. 263-286.
- [14] Millstein, R. (2016), “Probability in Biology: The Case of Fitness”, in Hájek and Hitchcock (eds.), *The Oxford Handbook of Probability and Philosophy*, Oxford University Press, pp. 601-622.
- [15] Pence, C. and G. Ramsey (2013), “A New Foundation for the Propensity Interpretation of Fitness”, *British Journal for the Philosophy of Science* 64, pp. 851-881.

- [16] Sober, E. (1984), *The Nature of Selection*, University of Chicago Press.
- [17] Sober, E. (2001), “The Two Faces of Fitness”, in R. Singh, D. Paul, C. Krimbas and J. Beatty (eds.), *Thinking about Evolution: Historical, Philosophical, and Political Perspectives*, Cambridge University Press, pp. 309-321.
- [18] Sober, E. (2010), “Evolutionary Theory and the Reality of Macro-Probabilities”, in E. Eells and J. H. Fetzer, eds., *The Place of Probability in Science, Boston Studies in the Philosophy of Science*, 84, pp. 133-161.
- [19] Sober, E. (2013), “Trait Fitness is not a Propensity, but Fitness Variation is”, *Studies in History and Philosophy of Biological and Biomedical Sciences*, 44, pp. 336-341.
- [20] Suárez, M. (2011) “Four Theses on Probabilities, Causes, Propensities” in *Probabilities, Causes and Propensities in Physics*, Springer, pp. 1-41.
- [21] Suárez, M. (2015), “Propensities, Probabilities and Experimental Statistics”, in M. Massimi et al. (eds.), *EPSA15 Selected Papers: European Studies in Philosophy of Science*, vol. 5, Springer, pp. 335-345.
- [22] Walsh, D., Ariew, A., and D. Mahen (2016), “Four Pillars of Statisticalism”, *Philos. Theor. Pract Biol*, 9 (1), pp. 1-18.

Do somatic cells really sacrifice themselves? Altruism and coercion in the evolution of multicellularity

Javier Suárez*
Jagiellonian University

Adrian Stencel**
Jagiellonian University

The evolution of multicellularity is one of those events in the history of life in which certain units lose their capacity to reproduce independently as a result of their formation of higher-level units capable of reproduction as a whole. In the case of the evolution of multicellularity, independently reproducing cells, as they evolved into multicellular individuals, lost their independent capacity for reproduction. Multicellular individuals have the capacity to reproduce themselves and to control the reproduction of the cells of which they are composed. These cells comprise two main subsets: somatic cells, whose reproductive regime is controlled and limited to the ontogeny of the multicellular organism, and germline cells, which give rise to cell lineages with the capacity of creating new multicellular individuals. The study of the evolutionary origin of multicellularity raises a question concerning the type of mechanisms that made this transition feasible.

One of the most pressing questions in the context of major transitions in evolution concerns the phenomenon of reproductive sacrifice. Why did certain formerly reproducing entities give up their reproductive potential to aid the reproductive abilities of other elements within a higher-level system? How do entities of this type evolve? As applied to the case of multicellular organisms, why do somatic cells sacrifice themselves in order to guarantee that germline cells can transmit their genes to the next generation, instead of doing the transmitting themselves? One conventional answer to this question refers to the theory of kin selection, or inclusive fitness (Bourke 2011; West et al. 2015). According to this explanation, since all the cells in a multicellular organism derive from a single cellular bottleneck, they are all nearly genetically identical, i.e., characterized by the highest possible degree of biological relatedness. This promotes altruistic cell-to-cell behavior, as the same genes will be found in the next generation regardless of which cell transmits

* jsuar3b@gmail.com

** adstencel@gmail.com

them. As a consequence, according to the kin selection explanation of multicellularity, the somatic cells of a multicellular organism sacrifice their reproductive capacity to enhance the reproduction of germline cells, since both subsets of cells bear the same (or very similar) genes. This explanation entails the consideration of multicellular organisms as altruistic or cooperative consortia of cells whose cohesion derives from the existence of altruistic cell-to-cell mechanisms, which have evolved due to a high degree of genetic relatedness (Michod and Roze 2001; Queller and Strassmann 2016; Folse and Roughgarden 2010).

The aim of this talk is to contest this idea by showing that the relevance of altruistic mechanisms in the evolution of multicellularity is more limited than advocates of kin selection theory suggest. We will show that contemporary research on multicellularity suggests that a stronger emphasis should be placed on so-called policing mechanisms (Clarke 2010, 2013), which are among the most important biological mechanisms making the evolution of multicellularity possible. Policing mechanisms involve biological structures that control any biological entity that compromises the survival of the organism in question, including multicellular organisms. In the case we study in this article, policing mechanisms involve the biological structures that actively control the excessive reproduction of somatic cells in order to guarantee the long-term survival of a multicellular organism and the transmission of germline cells. We will cite the pervasiveness of policing mechanisms in multicellular organisms to argue that multicellularity should be conceived as a system of coercion whereby germline cells guarantee their transmission to the next generation. Importantly, our view does not imply the lack of any role whatsoever for altruistic mechanisms in the evolution of multicellularity. The separation of germ and somatic cells has occurred many times, and in many different lineages, within the history of life; thus, we believe there is no single and unique way in which the evolution of multicellularity occurs. Furthermore, we believe that both altruism and coercion may have played a role in the evolution of multicellularity even within the same lineage, e.g., at different stages during its evolution. Thus, our point is that a more prominent role should be attributed to policing mechanisms in theoretical, experimental, and philosophical research on multicellularity than is currently the case due to the influence of the inclusive fitness view of multicellularity evolution.

Our talk will proceed in four stages. Firstly, we will briefly summarize an interpretation of the inclusive fitness theory of the origin of multicellularity, according to which somatic cells constitute altruistic consortia of cells that have evolved due to their genetic relatedness to germline cells. Then, we will show that the lack of reproduction in somatic cells may have two causal explanations. One explanation invites us to consider this lack a form of altruism; another, a form of coercion. Building on this, and relying on some

recent evidence from biological studies, we argue that it is better to think of somatic cells as being coerced to sacrifice themselves rather than as being altruistic. In this vein, we defend the idea that somatic cells do not sacrifice themselves voluntarily, but are rather coerced to do so. Multicellularity should thus be conceived as a coordinated system of coercion.

Perspectivismo científico

Margarita Vázquez*
Universidad de La Laguna

Manuel Liz**
Universidad de La Laguna

El perspectivismo científico es un tema importante y recurrente en filosofía de la ciencia. Entre los autores más relevantes cabe mencionar a Evandro Agazzi, Nancy Cartwright, Ronald Giere, Sandra Harding, Michela Massimi, Paul Teller o Bas van Fraassen. Este enfoque es sin duda interesante. Pero, de manera más o menos implícita, en todos ellos el perspectivismo tiene siempre un carácter local y aplicado. Y esto impide considerar al perspectivismo como una posición sustantiva, sobre todo frente a versiones moderadas del realismo y del relativismo. Las discusiones entre el realismo y el relativismo invisibilizan al perspectivismo.

Las discusiones entre el realismo y el relativismo han ocupado demasiado protagonismo. Y el perspectivismo ha acabado siendo algo muy similar a un realismo o un relativismo moderados. En nuestro trabajo, ofreceremos un perfil del perspectivismo capaz de hacer de él una posición filosófica sustantiva.

Hay una larga tradición del perspectivismo a veces cruzada con otras tradiciones. Entre los autores pertenecientes a ella estarían Leibniz, Nietzsche, Leibniz, Kant, Fichte, Nietzsche, William James, Brentano, Husserl, Russell, Wittgenstein, Hector-Neri Castañeda, Putnam o Dennett. Y también estarían los autores mencionados más arriba a propósito del perspectivismo científico. Y los rasgos que constituyen nuestro perfil serán sensibles a esa tradición. Después de presentar ese perfil, señalaremos algunas consecuencias importantes para el perspectivismo científico.

Un perfil del perspectivismo

Antes de dibujar un perfil para el perspectivismo, aclararemos brevemente las nociones de punto de vista y perspectiva, que utilizaremos de manera indistinta.

Adoptaremos un enfoque estructural. Un punto de vista tendrá la siguiente estructura canónica (Vázquez and Liz, eds., 2015):

* mvazquez@ull.edu.com

** manuliz@ull.es

$PV = \langle S, Act, n-CC, CC, Cp \rangle$, donde

1. S es el portador, o el poseedor, o el titular, del PV .
2. Act es un conjunto de relaciones que conectan a S con los contenidos explícitos del PV : $n-CC$ y CC . Cuando S es un sujeto personal, Act son actitudes psicológicas.
3. $n-CC$ y CC son las dos clases de contenidos que pueden estar incluidos explícitamente en el PV , respectivamente contenidos no-conceptuales y contenidos conceptuales. Los $n-CC$ son los contenidos cualitativos o fenomenológicos de la experiencia. Los CC son proposiciones conceptualmente articuladas.
4. Cp es un conjunto de condiciones de posesión para tener el PV .

Esta estructura permite identificar diversos constituyentes esenciales de las perspectivas y trazar una serie de distinciones importantes. Por ejemplo, las siguientes:

- D1 Contenidos explícitos de un PV y contenidos implícitos. Las Act de los sujetos personales pueden tener un $n-CC$ implícito. El sujeto es consciente de ellos y tienen condiciones de corrección.
- D2 Según sean los contenidos explícitos, el PV podrá ser no-conceptual, conceptual o híbrido.
- D3 Externo/interno a un PV y externo/interno al sujeto que adopta tal PV . Esta distinción crea un importante espacio de objetividad dentro de los PV .
- D4 Adoptar un PV es mantener una posición comprometida respecto a él. Pero también podemos mantener una posición distante si adoptamos un PV reflexivo respecto a él. Ninguna de estas posiciones implica a la otra.

Nuestro objetivo era elaborar un perfil para el perspectivismo. Proponer un perfil es algo muy distinto a definir. Y también algo distinto a hacer una reconstrucción racional. En criminología o diagnóstico médico, un perfil es adecuado en la medida que permita identificar aquello de lo cual el perfil es un perfil. También será así en nuestro caso.

Nuestro perfil del perspectivismo consta de cinco tesis y de tres ideas estratégicas. Las tesis darán sustantividad a la posición perspectivista, principalmente frente al realismo y al relativismo. Las tres ideas estratégicas son el marco general desde el que el perspectivismo hace frente a los problemas.

Cinco tesis perspectivistas

- T1 Dependencia perspectivista: Todos nuestros accesos a la realidad, y a nosotros mismos, dependen de nuestros PV en el sentido de que no podemos entender cómo podría existir esa realidad, ni nosotros mismos, aparte de la manera como se nos presentan desde los PV que adoptamos.
- T2 No igualitarismo: No todos los PV tienen el mismo valor.

- T3 Meliorismo: Siempre es posible adoptar *PV* mejores, es decir *PV* con más valor.
- T4 Invariantismo: Podemos reconocer la existencia de invarianzas relevantes acerca de las *Act*, con sus *n-CC* asociados, y acerca de los contenidos explícitos *n-CC* y *CC* que articulan nuestros *PV*.
- T5 External World: Como un caso particular de T4, a través de una gran variedad de puntos de vista se nos presenta un mundo de objetos y de sujetos teniendo propiedades, manteniendo relaciones y constituyendo hechos, eventos y procesos que no dependen de nosotros, que somos capaces de percibir de diferentes formas, y sobre los cuales podemos pensar y actuar.

T1 admite una lectura sólo epistemológica y también una lectura metafísica. La primera admite la posibilidad de que exista algo más allá de los *PV*, aun sin poder entender tal existencia. La segunda rechaza incluso esta posibilidad. Los perspectivismos locales y aplicados suelen adoptar una interpretación epistemológica. Esto es lo que habitualmente ocurre en el perspectivismo científico (Agazzi 2014, Giere 2006, Massimi 2016, Mitchell 2009,). Un perspectivismo generalizado y especulativo adoptaría la interpretación metafísica.

Consecuencias para el perspectivismo científico

Podemos obtener a continuación varias consecuencias importantes en relación al perspectivismo científico.

1. Sustantividad del perspectivismo

Esta es la consecuencia que más queremos destacar. La tesis T1, sobre todo en su interpretación metafísica, se opondría a todo realismo. Y esto incluye un realismo moderado que o bien limite ese realismo sólo a un dominio limitado, o bien considere que una parte de la realidad puede ser dependiente de las perspectivas y otra parte no.

El meliorismo de T3 (muy presente en Leibniz, William James y en Hilary Putnam), refuerza el distanciamiento respecto al realismo. Si todas nuestras perspectivas son mejorables, entonces deja de tener sentido suponer que alguna de ellas puede llegar a mostrarnos la realidad en sí misma, o la realidad al completo.

A su vez, la tesis T2 también se opondría al relativismo (Baghrarian y Coliva 2019, Boghossian 2006, Hautamäki 2020). Y T3, aún siendo independiente de T2, contribuye también a ese rechazo incluso en el caso de un igualitarismo local. A pesar de no implicar el rechazo del igualitarismo, T3 tiene un fuerte efecto disuasorio contra el igualitarismo. Intentar adoptar mejores perspectivas produce rápidamente desigualdades en su valor respectivo.

Esto mostraría la sustantividad del perspectivismo frente al realismo y al relativismo.

La tesis T1 puede favorecer el pluralismo. Pero debe subrayarse que se trata de un pluralismo no relativista. Y también lo hace la tesis meliorista T3, que sitúa al perspectivismo contra todo realismo y todo relativismo. T3 tiene una gran importancia en el perspectivismo. Actúa como motor de la dinámica de perspectivas. Adoptar una perspectiva y cambiar de perspectiva son procesos cuya orientación general es la búsqueda de perspectivas mejores.

2. Papel de los componentes personales en la ciencia

El origen de esos componentes personales serían siempre las *Act* involucradas en las perspectivas científicas en juego. Esas *Act* son indispensables en la ciencia.

Diferentes *Act* explican también planteamientos como la distinción de van Fraassen entre aceptación de la adecuación empírica de una teoría y creencia en su verdad.

3. Meliorismo científico

Respecto a la ciencia, el meliorismo de la tesis T3 implicaría algo así como un falsacionismo optimista.

Además de reforzar el rechazo del realismo y del igualitarismo relativista, ese meliorismo explicaría el compromiso con el mecanismo de la retractación en la dinámica de la ciencia (Massimi 2016).

Existiría un equilibrio inestable entre el mecanismos de retractación y el de reafirmación implicado por la fidelidad a la propia perspectiva.

4. Invariantismo

T4 (Nozick, 2003, Hales 1997) y T5 son compatibles con T1. El mundo externo sería una invarianza perspectivista. No es simplemente una proyección o construcción subjetiva, pero tampoco se sitúa más allá de las perspectivas. Se sitúa en ese espacio de objetividad que abre la distinción D3. Es algo interno a los PV, pero también algo externo a los sujetos que los adoptan. Y lo mismo se aplicaría a lo que podemos conocer de los propios sujetos.

En ese espacio de objetividad también se situarían toda verdad y racionalidad, en particular toda verdad y racionalidad científica.

Nos hemos fijado en el caso perspectivismo científico. Nuestra tesis de fondo es que el perspectivismo también sería una posición sustantiva en otros campos filosóficos.

Referencias

- [1] Agazzi, Evandro (2014). *Scientific Objectivity and its Contexts*. Heidelberg: Springer.
- [2] Boghossian, Paul (2006). *Fear of Knowledge. Against Relativism and Constructivism*. Oxford,:Oxford Univ. Press.
- [3] Baghramian, Maria & Annalisa Coliva (2019). *Relativism, New Problems of Philosophy*. London: Routledge.
- [4] Giere, Ronald (2006). *Scientific Perspectivism*. Chicago: Univ. of Chicago Press. Hales, Steven, 1997, “A Consistent Relativism”, *Mind*, vol. 196, 421: 33-52.
- [5] Hautamäki, Antti (2020). *Viewpoint Relativism. A New Approach to Epistemological Relativism based on the Concept of Points of View*. Heidelberg: Springer
- [6] Massimi, Michela (2016). “Four Kinds of Perspectival Truth”. *Philosophy and Phenomenological Research*, XCVI, 2: 342-359.
- [7] Michell, Sandra (2009) *Unsimple Truths: Science, Complexity, an Policy*. Chicago: The Univ. of Chicago press.
- [8] Nozick, Robert (2003) *Invariances. The Structure of the Objective World*, Boston: Harvard Univ.Press.
- [9] Vázquez, Margarita and Manuel Liz (eds.) (2015). *Temporal Points of View. Subjective and Objective Aspects*. Heidelberg: Springer.

La deuda técnica en el desarrollo de la inteligencia artificial y sus consecuencias

Enrique Villanueva*
Universidad de Oviedo

A pesar de que en sus orígenes, hace ya 65 años en la Conferencia de Dartmouth, la Inteligencia Artificial no contemplase en absoluto la introducción de las emociones en los sistemas que se estaban planteando desarrollar, manteniendo una visión puramente cognitivista de la que pronto se quejaron algunos grandes nombres como Hubert Dreyfus (1965), cuando Rosalind Picard (1997) acuñó el término “computación afectiva” a finales de los años 90, las emociones comenzaron a recibir mucha más atención a la hora de diseñar y conceptualizar los sistemas que se estaban creando.

Los avances en neurociencia y psicología han puesto de relieve la estrecha interconexión entre la corteza prefrontal, las áreas perceptivas del cerebro y los sistemas subcorticales relacionados con el procesamiento de las emociones (Rolls, 2018), de modo que el enfoque puramente cognitivista sobre la mente humana no parece que pueda ser útil para alcanzar el objetivo final de dotar de consciencia e inteligencia al estilo humano a sistemas artificiales.

David Levy (2008) escribió uno de los planteamientos pioneros sobre las futuras relaciones humano-máquina, explicando, con su optimismo característico, su convicción de que los robots emocionalmente competentes serán nuestros futuros amantes y compañeros sentimentales, una idea que ha encontrado gran aceptación desde entonces. Antes de estos planteamientos, las personas ya usaban máquinas para hacer sus experiencias sexuales más satisfactorias, e incluso mucho antes algunos proyectos en IA se comenzaron a usar con fines psicoterapéuticos, como el programa ELIZA a mediados de los años 60 (Bostrom, 2014: 7). Pese a que la inteligencia artificial ELIZA, por la tecnología disponible en su momento, no era capaz de establecer conversaciones muy convincentes, la mitad de los pacientes que la usaron en el hospital declararon que preferían hablar con ELIZA sobre sus problemas antes que con otro ser humano (Levy, 2008: 113). Este fenómeno ha dado lugar al llamado “Efecto Eliza”, que describe la tendencia a antropomorfizar las conductas de las inteligencias artificiales (Zhou y Fischer, 2019: 88).

La capacidad emocional de las máquinas, por tanto, es un elemento importante tanto para emular la inteligencia humana general como para desarrollar agentes que interactúen con las personas en contextos sociales o de cuidado. Las emociones, fingidas o sentidas,

* UO204495@uniovi.es

apoyan y dan consistencia a las acciones de las máquinas que las ejecutan (Pessoa, 2017: 819), además de aprovechar el efecto Eliza mencionado más arriba. El modelado de las emociones depende de la teoría que manejemos, y la cuestión de qué son las emociones y cómo funcionan está muy lejos de resolverse. La evidencia proporcionada por la fisiología, la neurología y la fenomenología ha dado lugar a una plétora de diferentes teorías que casan con los fenómenos observados, causando una infradeterminación difícilmente resoluble. Por tanto, la labor del ingeniero de arquitecturas emocionales consiste en muchos casos en seleccionar la que más se adapta a la naturaleza computacional –siendo la teoría OCC de Ortony, Clore y Collins (1988) o la de las emociones básicas de Ekman (1972) algunas de las más populares– y generar un modelo a partir de sus principios.

La emulación de los fenómenos observados sin tener una teoría sólida sobre la naturaleza de lo emulado no es un fenómeno nuevo en el campo de la computación, siendo otros ejemplos la producción y comprensión del lenguaje natural o la visión artificial (*computer vision*). La deuda técnica es un concepto usado en ingeniería de software para describir el coste adicional que debe ser pagado en el futuro como resultado de tomar un atajo (*shortcut*) cuando se desarrolla un sistema de software (Cristianini, 2019). Un ejemplo de esta deuda es el hecho de que muchos de los traductores en la web trabajen con muestras de traducciones ya realizadas por humanos, no entendiendo realmente lo que están traduciendo, lo que hará que si queremos máquinas que realmente entiendan la semántica del lenguaje tengamos que crear un método para esto más tarde (pagar la deuda).

La dimensión de esta deuda técnica adquiere un cariz social más relevante cuando pensamos, como se ha sugerido previamente, que estos sistemas de inteligencia artificial están siendo diseñados para ser nuestros terapeutas, amigos e incluso amantes. A juzgar por algunos fenómenos observados en la interacción humano-máquina, las personas tienden a responder bien y a adaptarse a las máquinas que simulan, aun torpemente, capacidades emocionales, incluso siendo conscientes de estar interactuando con un robot que no siente ni es autoconsciente (Nass y Moon, 2000: 94). Es el momento, pues, de reinterpretar esta deuda técnica con un enfoque centrado en sus consecuencias, como por ejemplo que los robots interactivos del futuro tendrán diferentes grados de precisión en la emulación de las emociones, teniendo un comportamiento adecuado en determinadas situaciones y fallando en otras, obligando al refinamiento paulatino hasta alcanzar un resultado satisfactorio, lo cual puede derivar en importantes consecuencias para pacientes o personas vulnerables que estén más en contacto con estos robots.

La plétora de preguntas abiertas que nos deja la emulación de emociones en sistemas artificiales incluye la privacidad de la información confidencial de los pacientes y usuarios y las imágenes sobre determinados grupos sociales, o de género, que aparecerán y se generalizarán. No solo estas cuestiones, sino también la necesidad de un grado mayor o menor de adaptación por parte de los usuarios que interactúen con los robots, entrando

dentro del alcance operativo de los sistemas, aceptando sus limitaciones para sacar más partido a sus ventajas; o el estatuto de las máquinas como sujetos morales, entre otros muchos.

La presente comunicación tendrá como objetivo aportar una visión crítica y sistemática de esta deuda técnica y sus consecuencias, con la intención contribuir a una visión más informada y crítica de la tecnología para que tanto los participantes en el proceso de diseño y producción de esta como los usuarios que la consuman puedan participar de forma más informada y democrática en el desarrollo de la tecnología que está por venir.

Bibliografía citada

- [1] Bostrom, N. (2014) *Superintelligence. Paths, Dangers, Strategies*. Oxford, Oxford.
- [2] Cristianini, N. (2019) “Shortcuts to Artificial Intelligence” en Pelillo, M. y Scantamburlo, T. (eds.), *Machines We Trust*. MIT Press, Cambridge.
- [3] Dreyfus, H. (1965) “Alchemy and Artificial Intelligence”, RAND Corporation, 1965. Recuperado de: <https://www.rand.org/pubs/papers/P3244.html>
- [4] Ekman, P., Wallace, F. y Ellsworth, P. (1972) *Emotion in the Human Face: Guidelines for Research and an Integration of Findings*, Pergamon Press, Nueva York.
- [5] Levy, D. (2008) *Love and Sex with Robots*, Cromwell Press Ltd, Wiltshire.
- [6] Nass, C. y Moon, Y. (2000) Machines and Mindlessness: Social Responses to Computers, *Journal of Social Issues*, 56, 1, 81-103.
- [7] Ortony, A., Clore, G. L y Collins, A. (1988) *The cognitive structure of emotions*, Cambridge University Press, Nueva York.
- [8] Pessoa, L. (2017) Do Intelligent Robots Need Emotion? *Trends in Cognitive Sciences*, 21, 11: 817-819.
- [9] Picard, R. (1997) *Affective Computing*, MIT Press, Cambridge.
- [10] Rolls, E. T. (2018), *The Brain, Emotion and Depression*, Oxford University Press, Nueva York.
- [11] Zhou, Y., Fisher, M. (ed.) (2019) *AI Love You. Developments in Human-Robot Intimate Relationships*, Springer, Nueva York.

Modeling Evolutionary Propensities

Cristina Villegas*
Universidad Complutense de Madrid

Evolutionary biology deals with probabilities at many different levels, from the stochasticity of genetic mutations to the probabilistic tendencies of populations to change in their composition by natural selection. Not only do the mathematical models of evolution make use of probabilistic tools, but the whole field incorporates a dispositional language (e.g., fitness, evolvability, robustness, modularity, plasticity) that suggests that chance is a central conceptual and methodological component for understanding and predicting evolution. Although philosophers have paid attention to this probabilistic aspect in some ways, new conceptual problems have arisen that demand a wider range of application of the tools of the philosophy of probability to evolutionary biology. In particular, these tools have only been partially incorporated into some aspects of the classical models of evolutionary genetics, mostly concerned with fitness differences, thus leaving aside the role of probabilities in the origin and impact of phenotypic variation. In this talk, I intend to fill a gap of this research by presenting some key similarities and differences between the evolutionary propensities of the classical framework of evolutionary genetics, such as fitness and the capacity to drift, and those of new methodological approaches interested in the nature and impact of variation, known as ‘variational tendencies’ (Wagner 2014), such as evolvability, variability, robustness, modularity and plasticity.

Variational tendencies used to serve as methodological assumptions in the classical framework of evolution rather than as a subject of study (Wagner and Draghi 2010), and only now is there an increasing interest in understanding the different tendencies of living systems to produce different degrees and types of variation. This has been part of philosophical debates regarding the theoretical and explanatory integration between the classical approach to evolutionary biology and more recent ones, what has come to be known as the Extended Evolutionary Synthesis (EES) (Pigliucci and Müller 2010, Huneman and Walsh 2017). Nonetheless, this inclusion of novel research approaches has left relatively untouched the mainstream philosophical understanding of chance, probabilities, and randomness in evolutionary biology. Within the classical picture of evolution, chance is usually associated with the lack of explanatory power of evolutionary factors, and it typically encompasses any factor other than natural selection, on the basis of their alleged ‘randomness’ (Millstein 2011). Any evolutionary result due to these factors, treated

* crvilleg@ucm.es

homogeneously, is considered a ‘contingent’ product of evolution. Natural selection, on the other hand, is considered to be the main driving force giving directionality to the evolutionary process. Within this classical framework, fitness has been understood for some decades now as a ‘propensity’, a term imported from the philosophy of probability in order to provide probabilities with a physical basis. This research has not dealt with the origin and impact of phenotypic variation, a current central concern beyond the classical models of evolutionary genetics. In the recent incorporation of different approaches to the study of evolution, nonetheless, the randomness of some non-selective evolutionary factors has been questioned, which suggests that randomness and probabilities are also notions to be revised within the debates over the EES.

Although variational tendencies have sometimes been alluded to in terms of propensities (e.g., Stoltzfus 2006), only very recently -and partially- have they been philosophically conceptualized in this way (Nuño de la Rosa and Villegas 2019). The very recent nature of these results means that the connections between the new understandings of the propensities of variation (e.g., evolvability) and the classical ones of ecological interactions (e.g., fitness) remain completely unexplored. In this talk, I want to present a tentative framework for the relationship between these novel ideas (i.e., variational tendencies as propensities) and the classical probabilistic ideas of evolutionary genetics (i.e., fitness and drift as propensities). This framework intends to encompass the interactions that hold in contemporary explanations and models of evolution between classical and variational propensities, an ingredient that shall become part of a broader philosophical picture of the explanatory structure of current evolutionary biology.

For doing this, I first present the most central evo-devo tool for modeling phenotypic variation, namely the genotype-phenotype (G-P) map. The G-P map represents an abstraction of how the genotypic differences between organisms give rise to differences in their phenotypes. The basic idea of a G-P map is that we can define a mapping function from specific genotypic variants to specific phenotypes, be it at a given population or at higher levels. The G-P map serves as a means for studying the properties of development that are relevant for the patterns of variation encountered in populations and higher taxa, both at the statistical level and at the level of causal explanations. Although they have been proven useful for scrutinizing patterns of extant variation (Hansen 2008) and specific developmental mechanisms responsible for them, G-P maps have also been very influential in the conceptualization of general variational tendencies as based in developmental properties. Similarly to how classical models of population dynamics make use of probabilistic properties –notably fitness– to predict and explain evolutionary changes, variational tendencies can be represented in G-P maps. For instance, a G-P map will typically contain different phenotypes that the system can develop into –such as brown, blue and green for eye color–, showing which phenotypes are more likely to arise under

genetic and environmental changes in reproduction, independently of their fitness value. In turn, these maps can represent properties such as the variability of a particular developmental system, serving as a basis for explaining and predicting patterns of phenotypic variation, and ruling out the idea that such variation is random in the classical sense.

Then, I show the main similarities and differences between classical and variational propensities. Among their similarities, I stress their analogous roles in causal explanations, as well as in establishing a probabilistic sample space in representational and predictive models. In particular, I argue that variational propensities entail a causal understanding of the developmental relation between genetic and phenotypic variation, analogous to fitness as a propensity entailing a causal understanding of the ecological capacities of individuals. Additionally, in abstracting away the causal structure relating genotypes and phenotypes, a G-P map sets a sample space of phenotypic possibilities that in turn can ground variational probabilities. This is similar to the role that fitness has in models of population dynamics as a mathematical expectation. Regarding their differences, these propensities are not conceived in the same way. While in the classical propensities of evolutionary genetics the causal-explanatory role is relatively decoupled from the mathematical-predictive one, this is not the case in evo-devo variational propensities. Particularly, I defend that not only statistical measures but comparative data, interventions and computational approaches are the bases of the propensities modeled in G-P maps, which are best understood as causal hypotheses.

Finally, evo-devo and classical population genetics notably represent different ways of providing explanations, the former alluding to typological one, while the latter makes use of populational ones. Although it has recently been argued that both views are most likely compatible (e.g., Lewens 2009), no integration has been made so far from the point of view of evolutionary models, predictions, or explanations. After introducing the main differences and similarities between the two types of propensities these two areas make use of, I present a tentative way in which these two types of explanations can be partially reconciled, by arguing that the idealizations involved in their derivations are very similar in short-term, microevolutionary models.

References

- [1] Hansen, T. F. (2008). Macroevolutionary quantitative genetics? A comment on Polly (2008). *Evolutionary Biology*, 35(3), 182-185.
- [2] Huneman, P., & Walsh, D. M. (Eds.). (2017). *Challenging the Modern Synthesis: adaptation, development, and inheritance*. Oxford University Press.
- [3] Lewens, T. (2009). What is wrong with typological thinking?. *Philosophy of Science*, 76(3), 355-371.

- [4] Millstein, R. L. (2011). Chances and causes in evolutionary biology: How many chances become one chance. *Causality in the Sciences*, 2, 425.
- [5] Nuño de la Rosa, L., & Villegas, C. (2019). Chances and Propensities in Evo-devo. *The British Journal for the Philosophy of Science*.
- [6] Pigliucci, M., & Müller, G. (2010). *Evolution—the extended synthesis*. The MIT Press.
- [7] Stoltzfus, A. (2006). Mutationism and the dual causation of evolutionary change. *Evolution & development*, 8(3), 304-317.
- [8] Wagner, G. P. (2014). *Homology, genes, and evolutionary innovation*. Princeton University Press.
- [9] Wagner, G. P., & Draghi, J. (2010). Evolution of evolvability. In Pigliucci and Müller (2010), pp. 379-399.

CIENCIA, TECNOLOGÍA Y SOCIEDAD

El arma del telescopio

María Dolores González Rodríguez*
Universidad de Salamanca

La realidad y la representación literaria

“Francisco de Quevedo es menos un hombre que una dilatada y compleja literatura”¹, esta cita resume la vastísima y compleja obra de Quevedo (1580-1645). Los escritos quevedescos denotan a un intelectual implicado en las zozobras políticas de la primera mitad del siglo XVII, conocedor de intrigas palaciegas y de asuntos exteriores (de primera mano por sus viajes, misiones diplomáticas y sus relaciones epistolares), pero también alguien capaz de diseccionar las cuestiones sociales más comunes. En sus escritos se encuentran ataques –en muchas ocasiones nada mesurados– a poderosos, a minorías, a costumbres y oficios, a los defectos humanos y las debilidades de todo tipo. La búsqueda de lo correcto y justo, a pesar de las contradicciones que puedan hallarse, parece ser una constante en la vida y obra de Quevedo. *La Hora de todos y la Fortuna con seso*² –en adelante *La Hora*– recorre esa diversidad sociopolítica y, en un momento, dedica su atención al telescopio y muestra su conocimiento sobre el avance del conocimiento científico.

La Hora es un escrito preparado en torno a los años 1633-1635, época en que Quevedo había trabajado en varios textos de tema político (*Marco Bruto*, la segunda parte de *La política de Dios*). *La Hora* fue terminada en marzo de 1636 –así consta en la dedicatoria del autor (véase nota 2)– pero, al igual que las otras obras políticas (a excepción de *Carta a Luis XIII*, 1635), no fue publicada en vida del autor, lo que lleva a pensar en una prevención de Quevedo ante las tensiones políticas del momento y la creciente animosidad con Gaspar de Guzmán, el conde-duque de Olivares. Finalmente, en 1639 el autor fue encarcelado hasta 1643, cuando cae en desgracia Olivares, figura muy presente en *La Hora* por su decisivo papel en los acontecimientos, medidas y decisiones políticas representadas.

Calificada como “fantasía moral”, *La Hora* recurre al tópico del “mundo al revés” para revisar satíricamente³ algunos de los males que acechan a la sociedad de su tiem-

* mdgr@usal.es

¹ “Quevedo” en *Otras Inquisiciones*, en Borges, 1996, 44.

² Empleamos *La Hora de todos y la Fortuna con seso* por ser la denominación que el autor expresó en la dedicatoria a Don Álvaro de Monsalve: “Su título es: *La hora de Todos, y la Fortuna con seso*» (QUEVEDO, 1987, 147); en el año 1650 fue publicada póstumamente la primera edición con el título *La Fortuna con seso y la hora de todos* (algunas ediciones actuales nombran así a la obra).

³ “El tratadillo, burla burlando, es de veras. Tiene cosas de las cosquillas, pues hace reír con enfado y desesperación. Extravagante reloj que dando una hora sola, no hay cosa que no señale con la mano.” (1987, p. 149).

po. La *Utopía*, su diseño de una sociedad justa, bien pudo influir en la disposición del tratado quevedesco; el escritor tuvo un ejemplar de la obra de Tomás Moro, recomendó su traducción y publicación (Jauralde, 1999: 746), además de verter al castellano y citar algunas partes (en *Carta a Luis XIII*, véase López Estrada, 1980, 84-92). Pero el aspecto utópico en el texto de Quevedo recae en los dioses del Olimpo, observadores que abren y cierran *La Hora*. El autor sitúa una reunión de los dioses para tratar las quejas por los males que asolaban la existencia humana; por imposición de Zeus la Fortuna –acompañada de la Ocasión– es obligada a conceder una hora para que “se hallen de repente todos los hombres con lo que cada uno merece” (Quevedo, 1987, 161). Tras esta apertura mitológica, el autor despliega un molde literario conformado por 40 “cuadros” –capítulos– en los que se enmarcan las diferentes escenas que componen la obra, desarrollando cada cuadro un tema que finaliza con la Hora en que las cosas tornan a la sensatez natural.

El telescopio y las Indias

El cuadro XXXVI, “Holandeses en Chile”, aborda el tema de las Indias y las preocupantes expediciones holandesas por tierras americanas⁴ (que Olivares ignoró). El encuentro entre un capitán holandés y el jefe araucano es descrito como una sibilina maniobra de ataque al Imperio español. El regalo de un telescopio –anteojo y también antojo– sirve al autor para caracterizar la estrategia de manipulación:

“El Holandés, conociendo la naturaleza de los Indios inclinada a juguetes y curiosidades, por engaitarlos la voluntad, los presentó [...] sombreros y espejos, y, últimamente, un *tubo* óptico, que llaman antojo de larga vista, encareciéndoles su uso, y con razón, diciendo que con él verían las naves que viniesen a diez y doce leguas de distancia, y conocerían por los trajes y banderas si eran de paz o de guerra, y lo propio en la tierra, añadiendo que con él verían en el cielo estrellas que jamás habían visto, y que sin él no podrían verse; que advertirían distintas y claras las manchas que en la cara de la Luna se mienten ojos y boca y en el cerco de Sol una mancha negra; y que obraba estas maravillas porque con aquellos dos vidrios traía a los ojos las cosas que estaban lejos y apartadas en infinita distancia”⁵ (Quevedo, 1987, 310-1).

Descritos los indígenas como ingenuos, confiados, frente a los sobornadores holandeses, el episodio gira al sorprenderles la Hora y el jefe araucano toma la palabra:

⁴ El cuadro se inicia con un tono propagandístico y crítico con Holanda y, en consecuencia, la defensa de España: «Dio una tormenta en un puerto de Chile con un navío de holandeses, que, por su sedición y robos, son propiamente dádiva de las borrascas y de los furiosos del viento» (Quevedo, 1987, 306-7).

⁵ Se han corregido las mayúsculas para ‘Sol’ y ‘Luna’. Las cursivas del original, *tubo* óptico, los editores señalan su corrección que en algunas ediciones figuraba la errata “cubo óptico”, nota 502).

“Instrumento que halla mancha en el Sol, y averigua mentiras en la Luna, y descubre lo que el cielo esconde, es instrumento revoltoso, es chisme de vidrio, y no puede ser bienquisto del cielo. Traer a sí lo que está lejos es sospechoso para los que estamos lejos; con él debisteis de vernos en esta grande distancia, y con él hemos visto nosotros la intención que vosotros retiráis tanto de vuestros ofrecimientos”⁶ (*ibid.*, 311-2).

Esta escena manifiesta el buen conocimiento de Quevedo sobre los hallazgos de Galileo publicados en *Sidereus Nuncius* (1610), además de características de los astros y las manchas de la Luna. Quevedo debió conocer otras investigaciones posteriores del italiano presentadas en la Academia Lincei en 1613, pues menciona “en el cerco del Sol una mancha negra”. Núñez Centella y Sánchez Ron (2010, 27) detallan cómo Galileo pudo observar las manchas solares y su publicación en 1613. El escritor no aporta conocimiento de los problemas de Galileo con la Inquisición, tampoco lo menciona.

El rechazo del telescopio, o de la novedad que representa, puede deberse a su procedencia, en línea con la misión propagandística del autor, más que a una consideración ideológica. El instrumento procede de un país protestante y enemigo de España, así todo el encuentro se transforma en una maniobra con fines políticos y territoriales.

El discurso del jefe indio en el momento en que llega la Hora resulta extraño en Quevedo dada su acendrada defensa de España, aquí por medio del personaje araucano reconoce la avaricia española y el sometimiento de los pobladores indígenas –también de holandeses–. El jefe no aboga por un frente común con los holandeses –idea que estos reforzaron el ofrecimiento del telescopio– sino que, en una especie de síndrome de Estocolmo, reconoce lealtad a la monarquía hispana. También afea el interés que mueve a los holandeses, rechazando el telescopio con el ruego:

“si nos queréis granjear, pues sois invencioneros, inventad instrumento que aparte muy lejos lo que tenemos cerca y delante de los ojos; que os damos palabra que con este que trae a los ojos lo que está lejos, no miraremos jamás a vuestra tierra ni a España” (*ibid.*, 313).

La conclusión del discurso es clara: “llevaos esta espía de vidrio, soplón del firmamento; que pues con los ojos en vosotros vemos más de lo que quisiéramos, no le habemos menester. Y agradézcale el sol que con él le hallastes la mancha negra, que si no, por el color intentárades acuarle y de planeta hacerle doblón” (*ibidem*). Como amarga conclusión quevedesca la avaricia mueve a los europeos, mientras el jefe araucano identifica el telescopio con un arma que les facilite la llegada a sus tierras.

⁶ Se han corregido las mayúsculas para ‘Sol’ y ‘Luna’.

Referencias bibliográficas

- [1] Borges, J.L. (1996). *Obras completas*, II. Buenos Aires, Emecé.
- [2] Galilei, Galileo (2010). *Noticiero sideral* (edición conmemorativa IV centenario *Sidereus Nuncius*, prólogo Núñez Centella, R. y Sánchez Ron, J.M.). La Coruña, Madrid, MUNCYT.
- [3] Jauralde Pou, P. (1999). *Francisco de Quevedo (1580-1645)*. Madrid, Castalia.
- [4] López Estrada, F. (1980). *Tomás Moro y España*. Madrid, Editorial de la Universidad Complutense.
- [5] Quevedo, F. (1987). *La Hora de todos y la Fortuna con seso* (edición de Jean Bourg, PierreDupont y Pierre Geneste). Madrid, Cátedra.

Dificultades en torno a la inclusión del libre albedrío como neuroderecho humano

José Manuel Muñoz*

Universidad de Navarra y Centro Internacional de Neurociencia y Ética (CINET) de la Fundación Tatiana

De acuerdo con el Potomac Institute for Policy Studies (2013, p. 11), podemos definirla neurotecnología como “cualquier tecnología empleada para investigar, modular, reparar o mejorar el sistema nervioso y su funcionamiento”. En los últimos años hemos asistido al crecimiento del debate acerca de cómo el uso de este tipo de tecnología puede no sólo permitir avances importantes en el ámbito de la salud neurológica y psiquiátrica, sino también afectar, en determinadas circunstancias, a derechos fundamentales de los ciudadanos. Un aspecto central de este debate tiene que ver con valorar si es conveniente incluir nuevos derechos humanos *ad hoc*, conocidos como “neuroderechos”, por considerarse que el actual marco normativo internacional resulta insuficiente para hacer frente a los retos ético-legales relacionados con la neurotecnología.

Actualmente existen dos propuestas de incorporación de neuroderechos. La primera (Iencay Andorno, 2017) incluye cuatro derechos: libertad cognitiva, privacidad mental, integridad mental y continuidad psicológica. La segunda, conocida como Iniciativa NeuroRights (Neurotechnology Center at Columbia University, n.d.; Yuste *et al.*, 2021; Yuste *et al.*, 2017), contempla cinco: identidad personal, libre albedrío, privacidad mental, acceso equitativo a la mejora mental y protección frente a sesgos algorítmicos. Cabe señalar que ya existen dos casos en los que se han adoptado neuroderechos a nivel estatal: la reforma constitucional chilena (La Tercera, 2021) y la Carta de Derechos Digitales española (Sánchez, 2020).

Si nos ceñimos específicamente al derecho al *libre albedrío*, comprobaremos que la Iniciativa NeuroRights lo define del siguiente modo: “Los individuos deberían tener *control* último [la cursiva es mía] sobre su propia toma de decisiones, sin manipulación desconocida procedente de neurotecnologías externas” (Neurotechnology Center at Columbia University, n.d.). Señalaré a continuación algunas dificultades que, a mi juicio, entrañaría incorporarlo como nuevo derecho (*cf.*: Herrera-Ferrá *et al.*, en preparación; Muñoz, 2019), ya sea mediante una eventual reforma de la Declaración Universal de Derechos Humanos (Naciones Unidas, 1948) o a través de ordenamientos jurídicos a escala nacional.

* jmunozorteg@unav.es

Dificultad n° 1.- Multidimensionalidad filosófica

La relación entre el libre albedrío y el determinismo (si ambos son compatibles o no) constituye un problema multidimensional e irresuelto en la tradición filosófica occidental. El espectro de posiciones que podemos encontrar es diverso: compatibilismo, determinismo duro, libertarismo, incompatibilismo duro, e incluso posiciones rupturistas que renuncian a la dicotomía compatibilismo-incompatibilismo (revisionismo, subjetivismo, autonomismo agnóstico) (Muñoz, 2012). Abogar por un derecho universal al libre albedrío debería pasar previamente por la búsqueda de una postura de consenso al respecto de este problema, tarea que se antoja larga y complicada si atendemos a la historia de la filosofía.

Dificultad n° 2.- Control último

Resulta llamativa la introducción del control último como elemento central de la definición de la Iniciativa NeuroRights. El control último de un agente, es decir, su capacidad para erigirse como autor genuino y fuente última de sus decisiones y/o acciones (Muñoz, 2012), es visto como una condición necesaria para que exista libre albedrío desde posiciones libertaristas, pero es criticada desde posiciones compatibilistas (e. g., Wolf, 2005, 261–262). Sería quizá conveniente que los responsables de la Iniciativa NeuroRights aclararan si incluir el control último en su definición de libre albedrío implica un respaldo a la concepción libertarista o, si no es así, por qué se ha introducido dicho término y a qué se refieren con él. También podría ser útil aclarar por qué en la versión en español de su web se ha sustituido el adjetivo “último” por “absoluto” y, en consecuencia, qué entienden por “control absoluto”.

Dificultad n° 3.- Diversidad cultural

Convendría prestar atención a la contextualización cultural en relación con el libre albedrío. Por ejemplo, un reciente estudio de Chernyak *et al.* (2019) con niños de Singapur y Estados Unidos reveló que los primeros “eran más propensos a profundizar sobre la falta de libre albedrío haciendo referencia al castigo y/o a tener que pedir permiso a las autoridades”, mientras que los segundos “eran más propensos a respaldar la libertad de actuar contra las normas” (Chernyak *et al.*, 2019, p. 866). Este tipo de estudios invitan a estudiar de qué modo cabe abordar una eventual implantación del derecho al libre albedrío prestando atención a la mayor priorización que a este, entendido como valor, parece otorgársele en las sociedades occidentales frente a las orientales.

Dificultad n° 4.- Conveniencia normativa

Finalmente, es importante preguntarse si la introducción del libre albedrío como nuevo derecho no resulta redundante a la luz de las disposiciones vigentes, con el consiguiente riesgo de que se produzca una sobrerregulación. En el caso de la Declaración Universal

de Derechos Humanos, podemos encontrar diversos artículos en los que la libertad para decidir o actuar aparece contemplada para distintas situaciones. He aquí algunos ejemplos (Naciones Unidas, 1948):

Art. 16.2: “Sólo mediante libre y pleno consentimiento de los futuros esposos podrá contraerse el matrimonio”.

Art. 18: “Toda persona tiene derecho a la libertad de pensamiento, de conciencia y de religión; este derecho incluye la libertad de cambiar de religión o de creencia, así como la libertad de manifestar su religión o su creencia, individual y colectivamente [...]”.

Art. 21.3: “La voluntad del pueblo es la base de la autoridad del poder público; esta voluntad se expresará mediante elecciones auténticas que habrán de celebrarse periódicamente, por sufragio universal e igual y por voto secreto u otro procedimiento equivalente que garantice la libertad del voto”.

Cabe añadir que el contenido de estos artículos parece ajustarse bien a una posición concreta en el problema del libre albedrío: el compatibilismo clásico, según el cual “*ser libre* significa poseer la capacidad para hacer lo que queremos, deseamos o escogemos hacer, y que no haya [ninguna coacción o constricción] que nos impida que ejerzamos dicha capacidad” (Muñoz, 2012, p. 8).

A mi modo de ver, las dificultades señaladas tienen gran calado y requerirían un debate amplio previo a cualquier procedimiento de implantación del libre albedrío como neuroderecho. Además, debe tenerse en cuenta el hecho de que proponer esta implantación supone apelar a una forma de autonomía personal de carácter ontológico, relativa a la autodeterminación de las decisiones y acciones, con las aristas metafísicas que eso conlleva y que no deberían ser ignoradas. En contraste, uno de los neuroderechos propuestos por Ienca y Andorno (2017) apela a la autonomía personal en forma de elección práctica, por lo que pertenece al ámbito de la ética más que al de la metafísica (Herrera-Ferrá *et al.*, en preparación): la libertad cognitiva, que podemos definir como “el derecho a modificar los estados mentales propios con la ayuda de neuroherramientas, así como a rehusar hacerlo” (Bublitz, 2013, p. 234). ¿Debemos considerar el libre albedrío y la libertad cognitiva como dos alternativas entre las que cabe escoger? ¿Se trata, en cambio, de derechos complementarios? Son cuestiones adicionales que escapan a los objetivos de este trabajo, pero sobre las que valdría la pena profundizar de aquí en adelante.

Referencias

- [1] Bublitz, J-C. (2013). My mind is mine!?! Cognitive liberty as a legal concept. En E. Hildt & A.G. Franke (Eds.), *Cognitive enhancement* (pp. 233–264). Springer. <https://doi.org/10.1007/978-94-007-6253-4>
- [2] Chernyak, N., Kang, C., y Kushnir, T. (2019). The cultural roots of free will beliefs: How Singaporean and U.S. children judge and explain possibilities for action in interpersonal contexts. *Developmental Psychology*, 55(4), 866–876. <https://doi.org/10.1037/dev0000670>
- [3] Herrera-Ferrá, K., Muñoz, J. M., Martínez, V., Nicolini, H., y Saruwatari, G. (En preparación). A contextual perspective on neurorights: Transnational and cultural considerations.
- [4] Ienca, M., y Andorno, R. (2017). Towards new human rights in the age of neuroscience and neurotechnology. *Life Sciences, Society and Policy*, 13, Artículo 5. <https://doi.org/10.1186/s40504-017-0050-1>
- [5] La Tercera. (2021, 13 de abril). *En histórica votación, aprueban proyecto de ley que regulará los neuroderechos en Chile*. La Tercera. <https://www.latercera.com/que-pasa/noticia/en-historica-votacion-aprueban-proyecto-del-ley-que-regulara-los-neuroderechos-en-chile/4IAQJIVHM5F75GRLAR2GQ27V24/>
- [6] Muñoz, J. M. (2019). Chile—Right to free will needs definition. *Nature*, 574(7780), 634. <https://doi.org/10.1038/d41586-019-03295-9>
- [7] Muñoz, J. M. (2012). Hacia una sistematización de la relación entre determinismo y libertad. *Daimon*, (56), 5–19. <https://revistas.um.es/daimon/article/view/141761>
- [8] Naciones Unidas. (1948). *Declaración Universal de Derechos Humanos*. Resolución 217 A (III), del 10 de diciembre, de la Asamblea General de las Naciones Unidas, París, Francia. <https://www.un.org/es/about-us/universal-declaration-of-human-rights>
- [9] Neurotechnology Center at Columbia University. (n.d.). *NeuroRights Initiative*. Columbia University. Recuperado el 15 de abril de 2021 de <https://nri.ntc.columbia.edu/>
- [10] Potomac Institute for Policy Studies. (2013). *Neurotechnology futures study*. Potomac Institute. <https://www.potomac institute.org/images/studies/NTFSReport-FINAL.pdf>
- [11] Sánchez, M. V. (2020, 17 de noviembre). *Los expertos opinan sobre la Carta de Derechos Digitales: “Un buen punto de partida muy mejorable”*. El País. <https://elpais.com/tecnologia/2020-11-17/los-expertos-opinan-sobre-la-carta-de-derechos-digitales-un-buen-punto-de-partida-muy-mejorable.html>

- [12] Wolf, S. (2005). Freedom within reason. En J. S. Taylor (Ed.), *Personal autonomy* (pp. 258-274). Cambridge University Press. <https://doi.org/10.1017/CBO9780511614194.012>
- [13] Yuste, R., Genser, J., y Herrmann, S. (2021). It's time for neuro-rights. *Horizons*, (18), 154-164. <https://www.cirsd.org/en/horizons/horizons-winter-2021-issue-no-18/its-time-for-neuro--rights>
- [14] Yuste, R., Goering, S., Agüera y Arcas, B., Bi, G., Carmena, J. M., Carter, A., Fins, J. J., Friesen, P., Gallant, J., Huggins, J. E., Illes, J., Kellmeyer, P., Klein, E., Marblestone, A., Mitchell, C., Parens, E., Pham, M., Rubel, A., Sadato, N., Specker Sullivan, L., Teicher, M., Wasserman, D., Wexler, A., Whittaker, M., y Wolpaw, J. (2017). Four ethical priorities for neurotechnologies and AI. *Nature*, 551(7679), 159-163. <https://doi.org/10.1038/551159a>

Valores contextuales para una ciencia mejor

Eulalia Pérez Sedeño*
IFS-CSIC

El estudio de la dimensión social de la ciencia muestra los efectos de la investigación científica sobre la vida humana y las relaciones sociales, pero también las consecuencias que las relaciones sociales y los valores tienen en la investigación científica. Tradicionalmente se ha considerado que el método científico elimina de la ciencia cualquier influjo subjetivo que pudiera entrar en ella. Es decir, impide que se incorporen valores supuestamente ajenos a la ciencia –políticos, religiosos, éticos...– restándole objetividad.

Pero, desde hace varias décadas, diversos análisis de distintas disciplinas, prácticas, o teorías científicas han mostrado la injerencia de valores contextuales o no cognitivos en la ciencia (por ejemplo, Bleier 1984; Haraway, 1989). El feminismo y los estudios de Ciencia, tecnología y género, por lo general, se han centrado en sacar a la luz prejuicios sexistas que interferían en lo que podríamos denominar “buena ciencia” (García Dauder y Pérez Sedeño 2017).

Sin embargo, la ciencia es una construcción social realizada por investigadoras e investigadores situados, es decir, que llevan a cabo su actividad en un tiempo y en un lugar, es decir situados en una cierta relación o relaciones con lo que se conoce y con otros sujetos cognoscentes. La ciencia es un proceso y una actividad de comunidades científicas insertas en contextos sociohistóricos concretos en cuyo seno encontramos valores personales, sociales y culturales, preferencias de grupo o individuales culturales, sociales, etc. que inciden o pueden incidir en diversos modos y grados sobre la práctica científica (Helen Longino, 1990; Haraway, 1989). De hecho, las y los investigadores pertenecen a diversas comunidades a la vez: la de su paradigma científico/intelectual, pero también a otras: políticas, ideológicas, religiosas, etc.

Pues bien, mostraré cómo los valores feministas pueden contribuir de manera positiva al desarrollo de una ciencia “mejor”. Tal es el caso de algunas investigadoras de finales del s. XIX y principios del XX descontentas con las ideas de la ciencia del momento que sustentaban la supuesta inferioridad social, física y mental de las mujeres (Gómez Rodríguez, 2004). Por ejemplo, muchos científicos, como Havelock Ellis, insistían en que la región frontal del cráneo –donde se hallaban los procesos abstractos– era relativamente mayor en los hombres, mientras que la parietal lo era en las mujeres. Pero, tras medir unos cuantos cerebros, se observó que la región frontal era relativamente mayor en las mujeres. Ellis (ibid. pp. 38-39)

* e.p.sedeno@csic.es

pasó a sostener que no había ninguna conexión especial entre tamaño de la región frontal y los procesos intelectuales y abstractos. Es decir, la evidencia disponible se interpretó de manera diferente para adecuarse a la creencia estereotipada. En palabras del propio Havelock Ellis, “En tanto las mujeres son diferentes en los caracteres sexuales primarios y la función reproductiva, nunca pueden ser absolutamente iguales, ni aun en los procesos psíquicos más elevados” (Ellis, 1894/1934: 405), pues “a las mujeres les desagradan los procesos esencialmente intelectuales de análisis o les desagradan también las normas y los principios rígidos y las proposiciones abstractas, entre otras” (Ellis, 1894/1934: 406).

Y este no es el único ejemplo, pues antropólogos o naturalistas como Charles Darwin, proclamaban que la mujer era un hombre que, ni física ni mentalmente, había evolucionado completamente (García Dauder y Pérez Sedeño, 2017: 22-27; Gómez Rodríguez, 2004: 61-95). Las teorías aceptadas de la época mantenían que las diferencias hombre-mujer se basaban en las características evolutivamente más beneficiosas para el hombre que para la mujer.

Pero, como señalaba al principio, muchas mujeres proto-feministas estaban descontentas con estas ideas y se aplicaron a rebatirlas. Tal fue el caso de Helen Thompson Woolley quien indicó que había otras explicaciones tan lógicas como las biológicas, a saber, las diferencias socio-educativas: “Las diferencias psicológicas del sexo parecen deberse en gran medida... a las diferencias en los influjos sociales que afectan al individuo en desarrollo” (Thompson Woolley, 1903: 14), citando como pruebas las diferencias en los juguetes, los juegos y el énfasis en la actividad física.

Esta autora realizó, en su tesis, la mayor investigación experimental de su época sobre diferencias sexuales, aplicando diversas pruebas a universitarias y universitarios para medir habilidades motoras, umbrales de sensibilidad, habilidades intelectuales y rasgos de personalidad. Sus resultados indicaban que no había diferencias en capacidades emocionales y que las diferencias intelectuales no eran significativas. También en este caso indicó que esas diferencias se debían a condicionantes sociales, no a factores biológicos innatos. Los resultados de su tesis aparecieron en dos libros publicados en 1903: *Mental Traits of Sex* y *Psychological Norms in Men and Women*.

El otro ejemplo se da en el siglo XXI, que ha visto resurgir una serie de prejuicios sobre las mujeres, en concreto en neurociencia, especialmente espoleada por los populismos de extrema derecha. En obras como las publicadas por Baron-Cohen (2003), Brizendine (2010) ahora no se dice que las mujeres sean inferiores en habilidades cognitivas y en todos los aspectos, como vimos antes, sino que las mujeres son diferentes a los hombres, por lo que las mujeres pueden decidir libremente qué actividades/estudios/profesiones desarrollar. La idea común es que hay un cerebro masculino y otro femenino y que, si se tiene un cerebro de uno u otro tipo, se pueden atribuir diferencias entre individuos en habilidades, logros, personalidad y expectativas.

Ante la persistencia de estas creencias y el resurgimiento de supuestos estudios científicos al respecto, nació a partir de 2010 lo que Sonia Reverter (2017) ha denominado la “guerrilla epistemológica” de la Neurogenderings Network, compuesta por un grupo de investigadoras que estudian cómo las complejidades de las normas sociales, las experiencias de vida variadas, los detalles de las condiciones de laboratorio y la biología interactúan para afectar los resultados de la investigación neurocientífica (<https://neurogenderings.wordpress.com/the-neurogenderings-network>).

Los estudios sobre sexo/género clasifican de manera sistemática a los participantes como mujeres u hombres y aplican procedimientos estadísticos para compararlos. Pero las diferencias son menos grandes de lo que se piensa. Hyde (2005) en su análisis de 46 estudios meta-analíticos sobre diferencias sexuales descubrió que prácticamente la mayoría de las puntuaciones de mujeres y hombres “se solapaban en la mayoría de las variables sociales, cognitivas y de personalidad” (citado en Rippon, Jordan-Young, Kaiser y Fine, 2014: 2). Por otro lado, ha sido difícil replicar informes que se daban por buenos y en los que se afirmaba que había diferencias de sexo/género en la organización funcional de las zonas del cerebro donde se localizan ciertas actividades cognitivas. Por tanto, se ha rechazado la idea de que hay dos tipos de cerebros claramente diferenciados: el masculino y el femenino (García Dauder y Pérez Sedeño, 2017: 33-61).

He escogido estos dos momentos porque son un ejemplo claro de cómo los valores intervienen –en este caso para bien– en la ciencia. El primero de los casos estudiados se produce en un momento en el que hay una tendencia científica muy fuerte que pretende justificar la inferioridad (intelectual, física y social) de las mujeres. Y es un momento también en el que, tras la Convención Seneca Falls de 1848, hay un fuerte movimiento reivindicativo, el sufragismo. Ante esa idea de la inferioridad de las mujeres, investigadoras que discrepaban de esas teorías y que pertenecían a comunidades epistémicas, pero también a la sufragista incorporan no sólo los valores de la comunidad psicológica de su época, abogando por métodos de investigación experimentales rigurosos, sino que también incorporan sus valores de igualdad y no inferioridad, lo que produce una nueva mirada sobre sus objetos de estudio y una reevaluación de la evidencia empírica. Lo mismo sucede en el segundo caso con las “nuevas” teorías neurosexistas que aparecen cuando el feminismo es imparable y se sirve de la “guerrilla epistemológica” en revistas y redes para desarrollar un nuevo neurofeminismo que incorpora los valores feministas, además de los de la comunidad neurocientífica.

No sólo la incorporación del feminismo como teoría crítica, ha tenido impacto en contenidos teóricos, prácticas y desarrollos científico-tecnológicos, sino que ha contribuido de manera significativa al desarrollo de una ciencia mejor, al introducir en la ciencia sus valores.

Referencias

- [1] Baron-Cohen, Simon (2003). *The Essential Difference: men, women and the extreme male brain*. Penguin/Basic Books.
- [2] Bleier, Ruth (1984). *Science and gender: A critique of biology and its theories on women*. Oxford: Pergamon Press.
- [3] Brizendine, Louann. (2010). *The male brain*. Londres: Random House.
- [4] Ellis, Havelock Henry (1894/1934). *Man and Woman*. Londres: Heinemann.
- [5] Fine, Cordelia. (2010/2011). *Cuestión de sexos*. Barcelona: Roca Editorial
- [6] García Dauder, Dau y Pérez Sedeño, Eulalia. (2017). *Las 'mentiras' científicas sobre las mujeres*. Madrid: La Catarata.
- [7] Gómez Rodríguez, Amparo (2004) *La estirpe maldita. La construcción científica de lo femenino*. Minerva Ediciones S. L, Madrid.
- [8] Haraway, Donna (1989): *Primate Visions. Gender, Race, and Nature in the World of Modern Science*, Routledge, Nueva York.
- [9] Longino, Helen E. (1990): *Science as Social Knowledge*, Princeton University Press, Princeton.
- [10] Reverter Bayón, Sonia (2017). “El Neurofeminismo frente a la investigación sobre la diferencia sexual”. *Daimon: Revista internacional de filosofía*, N° 6.
- [11] Rippon, Gina; Jordan-Young, Rebecca; Kaiser, Anelis; y Fine, Cordelia (2014). “Recommendations for sex/gender neuroimaging research: key principles and implications for research design, analysis, and interpretation”. *Frontiers on Human neurosciences*, vol. 8.

Percepción e implicación social en la regulación de la ciencia y la tecnología

Libia S. Requejo*
University of Salamanca

M. Isabel González-Bravo**
University of Salamanca

Introducción

La ciencia regulatoria, pese a la importancia social de las decisiones finales por su grado de afectación a los ciudadanos, se desarrolla en un contexto de controversias y falta de consenso, indeterminación e incertidumbre, deficiencias en los procesos de validación de hipótesis, o de urgencia con la que se necesita (Jasanoff, 1989; Rodríguez, 2004; Todt *et al.*, 2010). Adicionalmente, otros aspectos como la precaución, el bienestar social, los riesgos (Todt y Luján, 2011) o la presión de determinados grupos de interés con expectativas individualistas (Jasanoff, 1995), son tenidos en cuenta en los procesos, resultando cuestionable que el asesoramiento científico para la toma de decisiones pueda basarse, como se presume, en la precisión y legitimidad del proceso científico (Frewer and Salter, 2002).

¿Son los ciudadanos conscientes de esta multidimensionalidad existente en los procesos de reglamentación científico-tecnológicos? La ciencia regulatoria obtiene credibilidad en la medida en que es valorado el trabajo del científico, traduciéndose en una confianza en las decisiones políticas que de aquélla puedan derivarse.

Integrar a los ciudadanos en los procesos de regulación es una alternativa que se plantea como adecuada. Especialmente, como indican Bowman *et al.* (2015), en casos donde existe mayor riesgo percibido por las consecuencias de las alternativas que se están valorando o en épocas de crisis. Es ahí donde los valores de la sociedad pueden estar alineados con los valores de los reguladores, subsanándose así algunos conflictos existentes en varios ámbitos de la ciencia regulatoria.

El objetivo de este trabajo es estudiar el conjunto de criterios que, desde la perspectiva ciudadana, deben tenerse en cuenta en la regulación de la ciencia y la tecnología. En concreto, la confianza concedida por la ciudadanía al conocimiento científico como base

* libia@usal.es

** lola@usal.es

para la elaboración de leyes y regulaciones y la valoración de la participación ciudadana en esos procesos de regulación.

Metodología

El trabajo parte de los resultados de las Encuestas de Percepción Social de la Ciencia y la Tecnología (EPSCYT) realizadas por la Fundación Española para la ciencia y la Tecnología (FECYT) cada dos años. En concreto se han seleccionado cinco ítems que figuran en las encuestas desde 2004 hasta la actualidad (FECYT, 2004 a 2018):

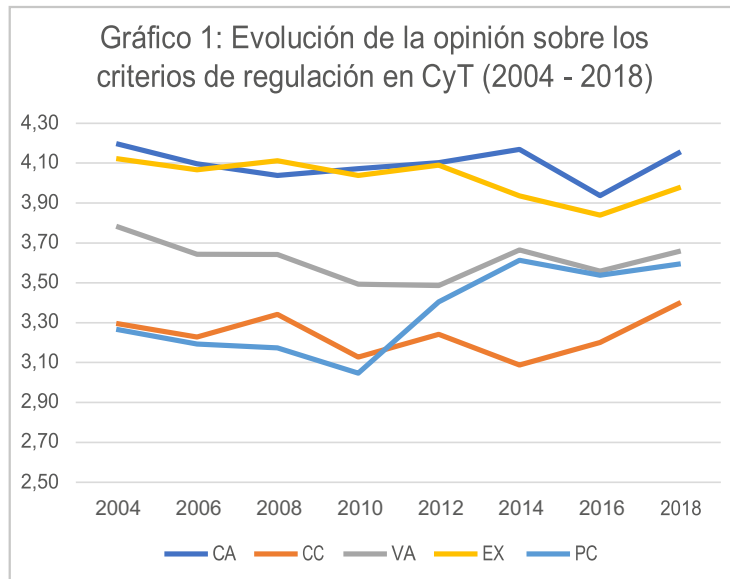
- Si no se conocen las consecuencias de una nueva tecnología, se debería actuar con cautela y controlar su uso para proteger la salud o al medio ambiente (CA).
- Los conocimientos científicos son la mejor base para elaborar leyes y regulaciones (CC).
- En la elaboración de leyes y regulaciones, los valores son tan importantes como los conocimientos científicos (VA).
- Las decisiones sobre asuntos de interés general relacionados con la ciencia y la tecnología es mejor dejarlas en manos de los expertos (EX).
- Los ciudadanos deberían tener un papel más importante en las decisiones sobre ciencia y tecnología que les afectan directamente (PC).

Adicionalmente se ha utilizado la información de una pregunta en la que los individuos hacen un balance general sobre la ciencia y la tecnología. Las tres opciones de respuesta utilizadas aquí posibilitan el establecimiento de perfiles de individuos y, con ellos, se palía en parte la necesidad de investigar el vínculo, sugerido por Todt (2011), de la confianza con la aceptación de la gobernanza y/o la participación:

- Los beneficios de la ciencia y la tecnología son mayores que sus perjuicios (B)
- Los beneficios y los perjuicios de la ciencia y la tecnología están equilibrados (E)
- Los perjuicios de la ciencia y la tecnología son mayores que los beneficios (P)

Resultados

La evolución de los valores medios poblacionales para cada atributo se recoge en el gráfico 1. En él se pone de manifiesto que todas las cuestiones analizadas en este documento deben tenerse en cuenta en la elaboración de las políticas públicas ya que la calificación media otorgada por los españoles supera el punto de indiferencia (que, dado que las opciones de respuesta fueron de 1 a 5, es 3) en todas ellas.



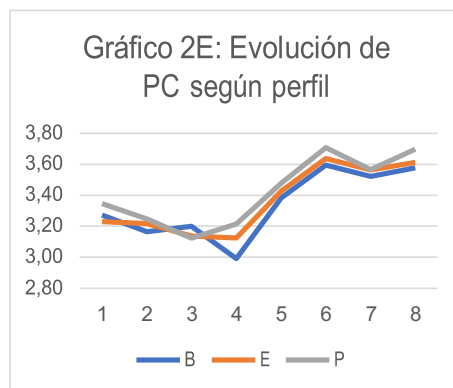
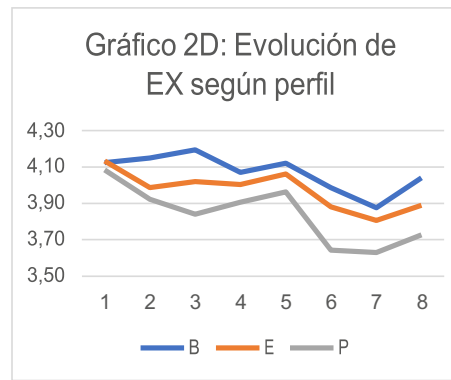
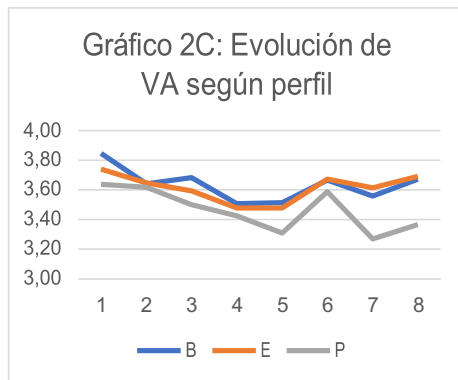
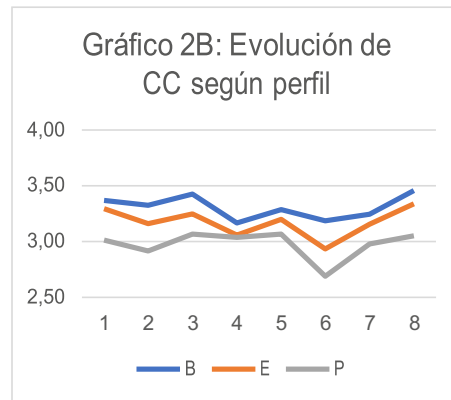
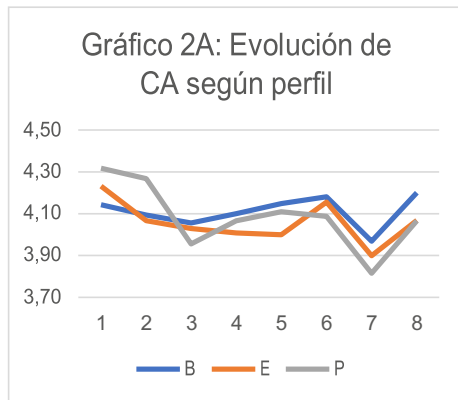
Fuente: elaboración propia a partir de FECYT (2004 a 2018)

No obstante, es destacable que para la sociedad española existen dos aspectos que vienen conviviendo en la opinión de la ciudadanía como prioritarios para la toma de decisiones sobre ciencia y tecnología, la cautela y la opinión de los expertos, si bien en los últimos años parece que ha disminuido la confianza depositada en la experticia frente a la cautela como principio de actuación en cuestiones tecnológicas vinculadas a la salud o al medio ambiente.

Por otra parte, la ciudadanía opina que los valores deben estar presentes en la formulación de políticas públicas y, en el periodo analizado, prefieren que aquéllos ocupen el mismo nivel que el conocimiento científico a que sea este la única base de las regulaciones.

Lo más llamativo es sin duda la evolución que se observa en la opinión sobre la participación ciudadana en decisiones que se tomen en el ámbito de la ciencia y la tecnología, rebasando desde hace casi una década el peso que debe atribuirse al conocimiento científico en cuestiones de regulación científico-tecnológica.

En los gráficos 2A a 2E se refleja la evolución experimentada por los criterios estudiados para cada uno de los colectivos formados según el balance que realizan de la ciencia y la tecnología. Asimismo, en la tabla 1 se resumen los resultados de la aplicación del Análisis de la Varianza y de las pruebas post hoc propuestas por Scheffé (1953) para determinar si existen diferencias significativas entre dichos colectivos.



Fuente: elaboración propia a partir de FECYT (2004 a 2018)

Tabla 1: Significatividad de las diferencias entre perfiles para cada criterio y año

| AÑO | CRITERIO | | | | |
|------|----------|-------|-------|-------|-------|
| | CA | CC | VA | EX | PC |
| 2004 | P>B | B=E>P | B>P | n.s. | n.s. |
| 2006 | P>B=E | B>E>P | n.s. | B>E=P | n.s. |
| 2008 | n.s. | B>E>P | B>P | B>E>P | n.s. |
| 2010 | B>E | B>E=P | n.s. | B=E>P | P=E>B |
| 2012 | B=P>E | B=E>P | B=E>P | B=E>P | n.s. |
| 2014 | n.s. | B>E>P | n.s. | B=E>P | n.s. |
| 2016 | B>P | B=E>P | B=E>P | B=E>P | n.s. |
| 2018 | B>E=P | B=E>P | B=E>P | B>P | n.s. |

p < 0,05

> perfil con valor superior.

= perfiles sin diferencia.

n.s. sin diferencias entre los tres perfiles.

Fuente: elaboración propia a partir de FECYT (2004 a 2018)

El principio de precaución ha sido valorado de manera diferente por los tres grupos estudiados en el periodo analizado, inicialmente eran aquellos que juzgan la ciencia como más peligrosa que beneficiosa quienes le daban un valor superior. En cambio, en los últimos años son los que ven más beneficios que perjuicios quienes valoran más la cautela como guía de acción en cuestiones científico-tecnológicas.

Destaca de nuevo de manera especial cómo valoran los diferentes perfiles la necesidad de participación ciudadana. La clara tendencia al alza expuesta antes se produce en igual sentido y magnitud en los tres colectivos. Es decir, la valoración del individuo sobre los beneficios o perjuicios de la ciencia es independiente de su opinión sobre la importancia de que la sociedad participe en las decisiones de ciencia y tecnología.

Conclusión

Los resultados demuestran que se han mantenido a lo largo de más de una década los hallazgos expuestos por Luján y Todt (2007) sobre la consideración de la precaución de los españoles como un valor establecido y de demanda social. Sin embargo, aquellos ciudadanos que confían más en los beneficios de la ciencia se muestran también a favor de una mayor cautela a la hora de tomar decisiones basadas en la ciencia. Más que una pérdida de confianza y legitimidad en la ciencia podríamos estar frente a la aceptación de que hay que tener en cuenta más variables a la hora de regular determinadas cuestiones en ciencia y tecnología con impacto social. Estos resultados serían coherentes con la defensa de una mayor participación de la ciudadanía dentro de los procesos de regulación por parte de este grupo como complemento al conocimiento científico.

En otro sentido, el crecimiento del apoyo a la participación que también se ha observado en el grupo que considera que la ciencia aporta más perjuicios podría descansar

más en una acción de control para que las decisiones no estén basadas únicamente en un conocimiento científico del que, en principio, parecen desconfiar.

Bibliografía

- [1] Bouwman, R., Bomhoff, M., de Jong, J.D., Robben, P. y Friele, R. (2015). The public's voice about healthcare quality regulation policies. A population-based survey. *BMC Health Serv Res* 15, 325. <https://doi.org/10.1186/s12913-015-0992-z>
- [2] FECYT (2004 a 2018). *Encuesta de Percepción Social de la Ciencia y la Tecnología en España*. FECYT. Madrid. <https://icono.fecyt.es/informes-y-publicaciones/percepcion-social-de-la-ciencia-y-la-tecnologia-en-espana> (consultado el 3 de mayo de 2021)
- [3] Frewer, L. and Salter, B. (2002). Public attitudes, scientific advice and the politics of regulatory policy: the case of BSE. *Science and Public Policy*, 29(2), 137-145.
- [4] Jasanoff, S. (1989). Norms for Evaluating Regulatory Science. *Risk Analysis*, 9(3), 271-273.
- [5] Jasanoff, S. (1995). Procedural Choices in Regulatory Science. *Technology in Society*. 17(3), 279- 293.
- [6] Luján, J.L., & Todt, O. (2007). Precaution in public: the social perception of the role of science and values in policy making. *Public Understanding of Science*, 16(1), 97-109.
- [7] Rodríguez, F.J. (2004). Normatividad en filosofía de la ciencia: el caso de la ciencia reguladora. *BiBLID*, 19(50), 173-190.
- [8] Scheffé, H. (1953). A method for judging all contrasts in the analysis of variance. *Biometrika*, 40(1-2), 87-110.
- [9] Todt, O. (2011). The limits of policy: Public acceptance and the reform of science and technology governance. *Technological Forecasting and Social Change*, 78(6), 902-909.
- [10] Todt, O. y Luján, J.L. (2011). Dinámica de los conceptos reguladores. Factores cognitivos y no cognitivos en el contexto de la toma de decisiones sobre los riesgos tecnológicos. *ÉNDOXA: Series Filosóficas*, 27, 317-336.
- [11] Todt, O., Rodríguez, J. y Fernández, I. (2010). Valores no epistémicos en la ciencia reguladora y en las políticas públicas de ciencia e innovación. *Argumentos de Razón Técnica*, 13, 41-56.

En la búsqueda de la visibilidad y el reconocimiento: la producción científica de las mujeres en el Perú

Alizon Rodríguez Navia*
Pontificia Universidad Católica del Perú -
Universidad Privada San Juan Bautista

I El problema

Han transcurrido 45 años desde la primera Conferencia Internacional de la Mujer auspiciada por la ONU y aunque los indicadores muestran un cierto progreso en cuanto a su participación en el campo de la CyT este no parece significativo. El porqué las cifras no revierten, pese a todos los esfuerzos desplegados en los últimos años, constituye una interrogante aún no resuelta.

El Perú como muchos países Latinoamericanos no escapa de dicha realidad. Si bien en los últimos años, el gobierno peruano a través del Consejo Nacional de Ciencia, Tecnología e Innovación Tecnológica – CONCYTEC ha desplegado un conjunto de esfuerzos significativos a fin de fortalecer las capacidades de investigación en CyT, aún se mantienen diferencias de género. De este modo, de acuerdo con datos correspondientes al 2020 –difundidos por el CONCYTEC de un total de 5028 investigadoras e investigadores a nivel nacional sólo el 31% son mujeres.

Según la UNESCO con base en la estadística de la Superintendencia Nacional de Educación Superior Universitaria (SUNEDU) , en el 2016 tres de cada diez egresadas de estudios de pregrado en la carreras de CyT o ingeniería eran mujeres. En el caso de maestrías dicha cifra se eleva a cuatro y en doctorado apenas llega a tres. Para el 2014 se estimó que del total de la población egresada de las universidades el 56.6% eran mujeres, pero su presencia en carreras de CTIM (ciencia, tecnología, ingeniería y matemática) representaba sólo el 32,9% (INEI, 2014). Del mismo modo, en torno al número de publicaciones en revistas indizadas en el Perú, de un total de 61,964 registros a julio 2016, la producción académica de las mujeres representa sólo un tercio del total.

La presente investigación toma como objeto de estudio a las ganadoras L’Oreal Perú, investigadoras de gran experiencia y trayectoria académica a fin de identificar, a partir de sus narrativas personales, sus avatares, búsquedas y disputas en el campo académico y los

* awrodriguez@pucp.edu.pe

sub-campos que la integran, (vida personal, vida académica y conciliación familia - vida personal) desde un enfoque de género.

II Los fundamentos teórico-metodológicos

Nuestra investigación es de orden cualitativo siendo nuestro abordaje teórico conceptual desde una perspectiva constructivista –estructural para lo cual recogemos los aportes del interaccionismo simbólico y la etnometodología (Flirck, 2012).

Realizamos un total de 24 entrevistas en profundidad de más de dos horas de duración cada una. Las entrevistas estuvieron dirigidas tanto a las 12 ganadoras L’Oreal (2010-2017) como a las 12 investigadoras/as del grupo de control que pertenecen a dos prestigiosas universidades del Perú (pública y privada). Para el ordenamiento de la información, hicimos uso del Atlas – Ti basado en la teoría fundamentada (TF).

El modelo de análisis de la presente investigación, asume el ámbito académico desde la noción de campo de Pierre Bourdieu (1988). Lo que implicó diseñar nuestra matriz de análisis, contemplando los campos de actuación (vida personal, vida académica y conciliación trabajo – vida personal) y también los campos de enunciación (interno, externo e interno-externo). Incluimos ambos elementos, en la medida que la actuación de las personas mantiene a nuestro juicio un orden bidimensional; es decir por un lado tenemos la actuación de los sujetos pero por otro lado tenemos el sentido de la actuación (Luckman y Berger, 1968). En síntesis, las personas actúan en función del sentido que otorga a los sucesos de la vida.

Los postulados teóricos se fundamentan en tres aspectos: los estudios Ciencia, Tecnología y Sociedad, los estudios sobre la producción académica de las mujeres, y finalmente la epistemología feminista (Harding, 1995).

III Los resultados

Los resultados que presentamos a continuación serán por cada uno de los sub campos que integran la vida académica, me refiero concretamente al subcampo de vida personal, conciliación familia-vida personal y el de la vida académica.

El campo de la vida personal, está conformado por las escenas más íntimas de la vida cotidiana las mismas que se encuentra cargada de discursos y mandatos normativos que modelan y pautan la vida de las mujeres, naturalizando no sólo las prácticas sexista sino también las valoraciones en torno al ejercicio académico de mujeres y hombres. En esa línea, no resulta extraño que aspectos como la maternidad, los roles en el hogar y la crianza de las y los hijos ocupen un lugar de prioridad en sus vidas a diferencia de los varones. Del mismo modo, características tales como; compromiso, minuciosidad, el mayor nivel de organización, la honestidad, la mayor sensibilidad etc son parte de las

características que según las ganadoras L'Oreal así como las investigadoras que pertenecen al grupo de control caracterizan el ser mujeres en la ciencia. Un aspecto que a nuestro juicio es relevante, es el rol que cumple el instinto en el proceso de investigación. Para las investigadoras, el instinto en la investigación es un rasgo que caracteriza la práctica femenina. Si bien, dicho recurso aparece como una capacidad de carácter complementario, constituye a nuestro juicio, un elemento revelador no sólo por los significados que las investigadoras otorgan a la misma sino también por vínculos que establece con la toma de decisiones y los sentidos que guían su acción.

En contraste, para los varones que pertenecen al grupo de control, las mujeres presentan mayor miedo al rechazo, esta afirmación constituye un elemento que comparten en común con las ganadoras L'Oreal así como con las investigadoras que pertenecen al grupo de control. Para las investigadoras, estar fuera del ojo público es mejor, ya que de este modo, se dedican de lleno a la labor de la investigación alejándose de toda situación de posible conflicto. La no visibilidad juega en este caso un rol perverso ya que por un lado es una estrategia de protección pero al mismo tiempo juega en contra de su empoderamiento.

En torno al campo de conciliación trabajo y vida personal. Se trata de un campo caracterizado por el sacrificio, la culpa, la presión y el estrés. La culpa y el sacrificio aparecen de manera más reiterada siendo el sacrificio de oportunidades, el más resaltantes. Asimismo, el sacrificio está vinculado a la importancia que ellas otorgan a los discursos de orden normativo que rigen la vida de las mujeres. Es por ello que utilizan este subcampo como espacio mediador, entre el subcampo de la vida personal y la vida académica, a fin de negociar y conciliar. Sin embargo, dicha tarea genera en ellas desgaste, sobre esfuerzo, presión y compromiso lo que a la larga afecta su salud mental y física.

Con respecto al campo de la vida académica, debemos señalar que se trata de un espacio de lucha y disputa del poder por el logro de la llamada "autoridad científica" lo que sin duda les otorga legitimidad académica (Bourdieu, 1983; Munévar, 2004). Sin embargo, es importante señalar que particularmente a las mujeres les resulta complejo acceder a dicha legitimidad, ello debido a que ésta se encuentra precedida del reconocimiento, el mismo que resulta esquivo más aún en áreas de CyT donde la presencia masculina resulta ser mayor. En ese sentido, si bien las ganadoras L'Oreal han logrado alcanzar un grado de autoridad académica en sus respectivas áreas, no ha sido una tarea fácil ni sencilla ya que muchas han vivenciado situaciones de incomprensión, discriminación y subvaloración de sus saberes. El campo académico por tanto se caracteriza como un campo donde prima la competencia y la presión, si bien dicha caracterización es altamente cuestionada, es finalmente aceptada y validada por la propia comunidad académica.

Otro aspecto de importancia dentro del mundo académico, tiene que ver con la noción de logro. El logro en este caso aparece asociado a la competencia, la productividad, el

financiamiento, la visibilidad y el reconocimiento. Sin embargo, en el caso de las mujeres, ello tiene sus bemoles ya que deben competir bajo reglas que han sido concebidas dentro una lógica masculina donde las tareas vinculadas a la maternidad y las tareas del cuidado no tienen cabida, lo que sin duda evidencia que las condiciones no son equitativas desde el partidor.

Finalmente, otro aspecto de importancia se vincula con la noción de discriminación. La discriminación en el campo académico cumpel el rol de obstaculizar el reconocimiento de los saberes de las mujeres y si bien no se presenta de manera explícita si de manera sutil, amparada en los reglamentos y sistemas de mérito y reconocimiento de las propias instituciones de educación superior.

IV Referencias bibliográficas

- [1] Berger, Peter, Luckmann, Thomas y Zuleta, Silvia (1968). *La construcción social de la realidad*. (975). Buenos Aires: Amorrortu.
- [2] Bourdieu, Pierre (1988a). *Homo academicus*. Stanford University Press.
- [3] Bourdieu, Pierre (1983). *Campo del poder y campo intelectual*. Buenos Aires: Folios
- [4] Flick, Uwe (2012). *Introducción a la investigación cualitativa* (Vol. 303). Ediciones Morata, SL.
- [5] Harding, Sandra (1995). *Feminismo y ciencia*. Barcelona: Ediciones Morata.
- [6] Instituto Nacional de Estadística (INEI) (2014). *II Censo Nacional Universitario 2010. Principales resultados*. Lima: INEI
- [7] Munévar, Dora Inés (2004). Construcción de conocimiento desde los márgenes. *Revista colombiana de Sociología*, (23), 181-215.

Anticipación e innovación responsable: los roles de las representaciones del futuro en la gobernanza científico-tecnológica

Sergio Urueña López*
University of the Basque Country, UPV/EHU

Desde principios de siglo y especialmente en el ámbito europeo, existe una tendencia institucional y académica a radicalizar las narrativas de responsabilización de las prácticas científico-tecnológicas y de innovación (CTI). Esta radicalización se propone tanto en términos extensivos (qué fases de los procesos CTI se consideran objeto de responsabilización) como intensivos (qué aspectos se tratan o consideran en relación a cada una de estas fases). Se trata de problematizar, bajo niveles crecientes de integración socio-técnica (Rodríguez, Fisher y Schuurbiens 2013), tanto los resultados y procesos de la investigación e innovación como los propósitos y expectativas a ellos asociadas. Todo ello convive además con una tendencia ‘inclusivista’ que reclama la necesidad de (i) alinear las prácticas de CTI con los intereses de diversos actores sociales (i.e., promover una CTI para la sociedad), así como (ii) hacer más partícipes a estos actores en los procesos de desarrollo científico-tecnológico (i.e., promover una CTI con la sociedad) desde etapas tempranas y a lo largo de todo el proceso de desarrollo (Owen, Macnaghten y Stilgoe 2012, European Commission 2013, 2017).

Sin embargo, estas tendencias hacia la politización de la CTI (Eizagirre, Rodríguez e Ibarra 2017), encuentran entre sus diversos desafíos dar respuesta al denominado dilema de Collingridge (1980): si bien los avances científico-tecnológicos y su direccionalidad son más susceptibles de ser modulados en sus etapas iniciales, en estas etapas solemos carecer de información y conocimiento. La reciente promoción de una CTI más responsable propone una problematización inclusiva de los procesos, resultados y propósitos desde las fases tempranas de la investigación y el desarrollo, pero esta propuesta normativa ha de lidiar con la realidad de que en estas primeras etapas el conocimiento suele ser un bien cualitativa y cuantitativamente insuficiente. La oportunidad de influir prematuramente en la direccionalidad de los desarrollos científico-tecnológicos se ve obstaculizada por las dificultades inherentes a cualquier búsqueda de orientación en contextos de ignorancia e incertidumbre; contextos donde los ubicuos problemas de legitimidad socio-epistémica son aún más acentuados (Grunwald 2011).

* sergio.urueña@ehu.eus

Las representaciones que apelan a futuros sociotécnicos son uno de los elementos que, si bien han reconocido como particularmente problemáticos (Nordmann 2013, 2014), entran en juego en los procesos de orientación y legitimación de la toma de decisiones del presente (Poli y Valerio 2019, Poli 2019). Estos futuros sociotécnicos apelan simultáneamente a la coevolución de potenciales desarrollos científico-tecnológicos y a (sus) potenciales efectos y cambios sociales (Konrad y Böhle 2019, 101) y son objeto de controversias y disputas socio-políticas. Esto último es así no sólo en tanto apuntan a eventos que carecen de existencia y, por lo tanto, se formulan típicamente sobre la base de supuestos que presentan una alta precariedad epistémica, sino también en tanto portan consigo valores, intereses y otra serie de elementos normativos (e.g., las nanotecnologías, la inteligencia artificial o la biotecnología) (Arnaldi 2014, Grunwald 2010). Los futuros sociotécnicos expresan, y consolidan, dinámicas epistémico-políticas que configuran la coproducción de proyectos y líneas CTI emergentes (Jasanoff 2020).

El desconocimiento, incertidumbre y multiplicidad de intereses en juego no ha servido, sin embargo, para disuadir a un amplio abanico de propuestas normativas –alineadas con las tendencias ‘inclusivistas’ señaladas anteriormente– de hacer uso de los “futuros sociotécnicos” como un instrumento metodológico-interventivo (Lösch et al. 2019, Konrad et al. 2016). Ejemplos de ello son propuestas como Future-Oriented Technology Assessment (Nazarko 2017), Real-Time Technology Assessment (Guston y Sarewitz 2002), Technology Assessment (Grunwald 2019), Anticipatory Ethics (Brey 2012), Anticipatory Governance (Barben et al. 2008), Responsible Innovation (Stilgoe, Owen y Macnaghten 2013, Owen y Pansera 2019) y Responsible Research and Innovation (von Schomberg 2013, European Commission 2013). Todas estas propuestas incluyen explícitamente a la anticipación (o a herramientas anticipatorias tales como ‘foresight’) entre sus pilares operativos centrales.

Aunque la anticipación es reconocida en la actualidad por numerosas propuestas y marcos normativos como una herramienta clave para la responsabilización de la CTI, la manera en que esta es comprendida y operacionalizada es variable. Por un lado, los supuestos ontológicos, epistemológicos y normativos que operan en cada uno de los marcos normativos conforman diferentes formas de entender y operacionalizar la anticipación. Por otro, cada modelo normativo aplica la anticipación a diferentes dominios, centrándose en diferentes modular diferentes roles que de facto ejercen las representaciones del futuro en la gobernanza científico-tecnológica.

La presente comunicación sostendrá que la finalidad y el alcance de la anticipación como herramienta socio-epistémica orientada a la responsabilización de la CTI únicamente pueden comprenderse en toda su radicalidad y amplitud teórico-conceptual si previamente se reconocen los distintos roles que desempeñan las representaciones del futuro en las dinámicas de gobernanza de la investigación y la innovación. En concreto, y

partiendo del estado de la cuestión esbozado en los párrafos anteriores, se defenderá que la anticipación adquiere sentido y significancia en tanto herramienta instrumental-interventiva en relación a tres roles generales que las representaciones del futuro juegan en las dinámicas de ciencia, tecnología e innovación.

Referencias

- [1] Arnaldi, Simone. (2014). “Unlocking the Futures of Nanotechnology. Future-Oriented Narratives and Access to the Public Discourse on Nanoscale.” En *In Pursuit of Nanoethics*, editado por Bert Gordijn y Anthony Mark Cutter, 183-192. Dordrecht: Springer.
- [2] Barben, Daniel, Erik Fisher, Cynthia Selin, y David H. Guston. (2008). “Anticipatory Governance of Nanotechnology: Foresight, Engagement, and Integration.” En *The Handbook of Science and Technology Studies*. Third Edition, editado por Edward J. Hackett, Olga Amsterdamska, Michael Lynch y Judy Wajcman, 979-1000. Cambridge, MA and London: The MIT Press.
- [3] Brey, Philip A. E. (2012). “Anticipatory Ethics for Emerging Technologies.” *NanoEthics* 6 (1):1-13. doi: 10.1007/s11569-012-0141-7.
- [4] Collingridge, David. (1980). *The Social Control of Technology*. London: Francis Pinter Ltd.
- [5] Eizagirre, Andoni, Hannot Rodríguez, y Andoni Ibarra. (2017). “Politicizing Responsible Innovation: Responsibility as Inclusive Governance.” *International Journal of Innovation Studies* 1 (1):20-36. doi: 10.3724/sp.J.1440.101003.
- [6] European Commission. (2013). Horizon 2020, Work Programme 2014-2015: 16. *Science with and for Society*, C(2013) 8631 of 10 December 2013. Brussels.
- [7] European Commission. (2017). Horizon 2020, Work Programme 2016-2017: 16. *Science with and for Society*, C(2017)2468 of 24 April 2017. Brussels.
- [8] Grunwald, Armin. (2010). “From Speculative Nanoethics to Explorative Philosophy of Nanotechnology.” *NanoEthics* 4 (2):91-101. doi: 10.1007/s11569-010-0088-5.
- [9] Grunwald, Armin. (2011). “Responsible Innovation: Bringing together Technology Assessment, Applied Ethics, and STS research.” *Enterprise and Work Innovation Studies* 7:9-31.
- [10] Grunwald, Armin. (2019). “The inherently democratic nature of technology assessment.” *Science and Public Policy* 46 (5):702-709. doi: 10.1093/scipol/scz023.

- [11] Guston, David H., y Daniel Sarewitz. (2002). "Real-time technology assessment." *Technology in Society* 24 (1-2):93-109. doi: 10.1016/s0160-791x(01)00047-1.
- [12] Jasanoff, Sheila. (2020). "Imagined worlds: The politics of future-making in the twenty-first century." En *The Politics and Science of Prevision: Governing and Probing the Future*, editado por Andreas Wenger, Ursula Jasper y Myriam Dunn Cavelty, 27-44. London: Routledge.
- [13] Konrad, Kornelia, y Knud Böhle. (2019). "Socio-technical futures and the governance of innovation processes –An introduction to the special issue." *Futures* 109:101-107. doi: 10.1016/j.futures.2019.03.003.
- [14] Konrad, Kornelia Elke, Harro van Lente, Christopher Groves, y Cynthia Selin. (2016). "Performing and Governing the Future in Science and Technology." En *The Handbook of Science and Technology Studies*, Fourth Edition, editado por Ulrike Felt, Rayvon Fouche, Clark A. Miller y Laurel Smith-Doerry, 465-493. Cambridge, MA: MIT Press.
- [15] Lösch, Andreas, Armin Grunwald, Martin Meister, y Ingo Schulz-Schaeffer. (2019). *Socio-Technical Futures Shaping the Present. Empirical Examples and Analytical Challenges*. Wiesbaden: Springer.
- [16] Nazarko, Łukasz. (2017). "Future-Oriented Technology Assessment." *Procedia Engineering* 182:504-509. doi: 10.1016/j.proeng.2017.03.144.
- [17] Nordmann, Alfred. (2013). "(Im)Plausibility²." *International Journal of Foresight and Innovation Policy* 9 (2/3/4):125-132. doi: 10.1504/ijfip.2013.058612.
- [18] Nordmann, Alfred. (2014). "Responsible innovation, the art and craft of anticipation." *Journal of Responsible Innovation* 1 (1):87-98. doi: 10.1080/23299460.2014.882064.
- [19] Owen, Richard, Phil Macnaghten, y Jack Stilgoe. (2012). "Responsible research and innovation: From science in society to science for society, with society." *Science and Public Policy* 39 (6):751-760. doi: 10.1093/scipol/scs093.
- [20] Owen, Richard, y Mario Pansera. (2019). "Responsible Innovation and Responsible Research and Innovation." In *Handbook on Science and Public Policy*, editado por Dagmar Simon, Stefan Kuhlmann, Julia Stamm and Weert Canzler, 26-48. Cheltenham and Northampton, MA: Edward Elgar.
- [21] Poli, Roberto. (2019). *Working with the Future: Ideas and Tools to Govern Uncertainty*. Milano: Bocconi University Press.
- [22] Poli, Roberto, y Marco Valerio. (2019). *Anticipation, Agency and Complexity*. Cham: Springer.

- [23] Rodríguez, Hannot, Erik Fisher, y Daan Schuurbiens. (2013). “Integrating science and society in European Framework Programmes: Trends in project-level solicitations.” *Research Policy* 42 (5):1126-1137. doi: 10.1016/j.respol.2013.02.006.
- [24] Stilgoe, Jack, Richard Owen, y Phil Macnaghten. (2013). “Developing a framework for responsible innovation.” *Research Policy* 42 (9):1568-1580. doi: 10.1016/j.respol.2013.05.008.
- [25] von Schomberg, René. (2013). “A Vision of Responsible Research and Innovation.” En *Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society*, editado por Richard Owen, John R. Bessant y Maggy Heintz, 51-74. Chichester: Wiley.

SYMPOSIUM:
ONTOLOGICAL AND METAONTOLOGICAL
DEFLATIONISMS

A Justification of the relevance of the symposium

This symposium is part of a research project that seeks to go deeper into the analysis and formulation of some of the main proposals considered deflationary and / or minimalist in ontology and meta-ontology. Metaontological deflationism is to be understood as related to (although different from, and indeed logically independent from) ontological deflationism, relative to the notion of existence. Deflationism at the metaontological level supposes a higher level of abstraction than mere ontological deflationism since it is a form of deflationism about the character or nature of ontology itself, ontological disputes, and metaphysical controversies about existence. The debate around these questions is a very rich and alive one at the moment. The symposium contributes by addressing some of the questions involved in this debate and consists of five talks. The first talk deals with the elucidation of the deflationary nature of different philosophical views that, while being different and committing to various thesis, have nevertheless been called “neo-Carnapian”; moreover, it determines how each view is Carnapian or not. The symposium continues to analyze deflationism by studying how deflationist views about different concepts relate to each other. One of them provides a counterexample to Thomasson’s (2014) proposal that semantic deflationism entails metaontological deflationism and aims to help deflationism delineate more clearly their adversary and their own position. The other analyzes Rayo’s (2013) view to see about which notions he can be considered a deflationist, and to assess how his views relate to deflationism about different concepts. Finally, the two remaining talks discuss two semantic deflationist readings respectively of fictionalist discourse and indexicals. Deflationary realist accounts of fictional characters are discussed and rejected given there is no reasonable way of deciding among them; an alternative non-realist account is defended that allows us to disregard the issue. Finally, the ontological implications of McTaggart’s and Perry’s accounts of the indexicals’ now’ and ‘I’ are discussed. The author argues that explaining the role these indexicals play does not require postulating particular ontological objects, properties, events, or relations. Alleged ontological issues derived from the study of indexicals are claimed to be nonsubstantive, in the sense that nothing ontologically substantial is at stake.

The contributions will represent an advance in the current discussion on different forms of deflationism in terms of conceptual analysis of metaontological proposals and the ontological implications of deflationary semantics. The convergence of these talks under the umbrella provided by the symposium concerning metaontological and ontological deflationism is meant to allow the authors to set in common their insights in various arenas to clarify and advance the discussion.

Abstract for the symposium

The symposium includes five contributions. The first examines several so-called neo-Carnapian metaontological and ontological views to clarify the notion of metaontological deflationism and to argue that none of them provides an adequate characterization of Carnap's view. The second and third contributions deal with the relations among different sorts of deflationisms. One of them illustrates that there is no implication relation between semantic and existence deflationism and metaontological deflationism. The other characterizes and explores how deflationism about objects and existence relates to metaontological deflationism in Rayo's (2013) proposal. Finally, the two remaining contributions discuss deflationist views about particular cases: fictional entities and indexicals. The first one sets a contrast between deflationary realist accounts of fictional discourse and an alternative fictionalist proposal, to finally outline the reasons to prefer the latter. The last contribution contends that accounting for the role certain indexicals play does not require postulating particular ontological objects, properties, events, or relations.

Carnap's metaontological deflationism and some neo-Carnapian approaches

Xavier De Donato-Rodríguez*
University of Santiago de Compostela

José L. Falguera**
University of Santiago de Compostela

In recent times, there has been a significant revival of Carnap's ideas on metaontology. As Berto and Plebani (2015) recall, Quine's criticism against the analytical/synthetic distinction was a motivation for many to ignore or reject Carnap's ideas on ontological questions. On the other hand, Thomasson (2015, 29) argues that Carnap's original ideas were dismissed because they were wrongly associated with antirealism and verificationism. If it may be true that Carnap's views lost popularity after Quine's critique, it is also clear that, perhaps leaving aside other related theses that have been associated with Carnap and logical positivism, Carnap's approach to ontological questions has been revived in recent years, to the extent that different authors have been called neo-Carnapians for defending similar views. A significant part of the debate in contemporary ontology and metaontology in analytic philosophy has to do with the so-called discussion between neo-Quineans and neo-Carnapians, with the former dominating the philosophical scene for several years and the latter gaining prominence more recently. As Berto and Plebani put it, "[m]any contemporary philosophers working on metaontology are grouped under the label of *deflationists*, meaning that, for them, many or most ontological questions are somehow shallow or unsubstantial and we should devalue their importance [...] Carnap is often cited as a source of inspiration for such philosophers." (Berto and Plebani 2015, 72-73).

Among the philosophers that have been called 'neo-Carnapians' or have been associated with Carnap's approach to ontological questions, we may count very different authors that, though somehow related to Carnap (1950), are holding very different theses or, in any case, express similar ideas in very different ways. Some of the approaches that have been called 'neo-Carnapian', or are somehow related to Carnap, are Rayo's (2013) 'trivialist Platonism', Thomasson's (2015) easy-approach, Hirsch's (2011) quantifier variantism, Schiffer's (2003) pleonastic account, and Putnam's (1987) conceptual relativity,

* xavier_donato@yahoo.com

** joseluis.falguera@usc.es

among other proposals. Neither these ways are equivalent, nor all these theses can be legitimately attributed to Carnap. Furthermore, in some cases, these theses are such that they should be carefully distinguished and acknowledged as opposed. The present contribution is part of a more ambitious project that aims to examine to what extent neo-Carnapians are indebted to Carnap's original ideas or are really faithful to him. In this particular case, we will focus on the thesis of quantifier variantism (see Hirsch 2002, 2009, 2011; and Putnam 1987) as well as on Hofweber's (2009, 2016) interpretation of Carnap in terms of the double reading of the existential quantifier, and we will argue that neither of both approaches legitimately capture Carnap's own kind of ontological deflationism.

According to quantifier variantists, the meaning of the existential quantifier and of the verb 'exist' can vary depending on the language or conceptual scheme we use. It is not only that there is no privileged unique language that may represent reality, but it is also that the meaning of 'exists' may vary depending on our linguistic choices. This is also a form of ontological deflationism, as ontological debates are usually rejected for being not properly substantial, but rather verbal disputes (see Hirsch 2011, chap. 9). Many authors have identified Hirsch's and Putnam's quantifier variantism as a genuinely Carnapian form of deflationism and understand Carnap himself as defending this view (see, for instance, Sider 2001, Fine 2009, and Eklund 2009, among others). However, the fact is that this was not Carnap's position, and in no way can it be considered Carnapian. Concerning this, we agree with Thomasson (2015, chap. 1), who warns us that the idea that the existential quantifier, or the word 'exists', may vary in their meaning has nothing to do with Carnap. Only one sense of the existential quantifier has relevance, as the notion of existence and of the existential quantifier is a univocal, formal concept.

Hirsch--while considering himself a neo-Carnapian- takes distance from Carnap when says: "I have a problem, however, in calling Carnap a quantifier variantist, insofar as he is often viewed as a verificationist and anti-realist" (Hirsch 2011, xvi). This is because Hirsch understands quantifier variantism as fully compatible with realism. According to this, Carnap would not count as a quantifier invariantist. However, Hirsch's reasons are the wrong reasons. Carnap is not a quantifier variantist not because he was an anti-realist, in the sense of Putnam or in any other sense. Nor because he was a verificationist. When Carnap wrote his seminal paper on ontology, he had already abandoned the verificationist theory of meaning. On the other hand, to understand Carnap correctly, we must consider that his position cannot be confused either with, of course, any form of ontological realism or with any form of antirealism. Carnap's position is equidistant from both and can be called 'agnostic.' According to Carnap, only internal ontological questions are meaningful (e.g., have cognitive value, logical or empirical), though they receive a straightforward answer within a linguistic framework. External questions, instead, go far beyond linguistic frameworks and are meaningless. The existential quantifiers carry

ontological commitment but just internally to the framework. Any ontological commitment in the external sense is rejected. Ontological debates, for Carnap, are thus deflated in the sense that, if we persist in asking whether, far beyond their framework, numbers or propositions actually exist, the controversy will continue for centuries because there cannot be an answer (to the external question). From a purely internal sense, ontological questions are perfectly fine and receive a straightforward answer by logical or empirical means following the rules of the framework (in this respect, Carnap's view certainly has similarities with Thomasson's).

Finally, we will focus on whether Carnap is better to be understood in deflationary terms, or in any other different way, as Hofweber (2016) and Kraut (2016) have argued. We will see that, no matter whether we take Hofweber's line of making the existential quantifier polysemic, or if we follow Kraut's expressivist proposal, we equally separate from Carnap's position in order to embrace very different theses.

Truth is not enough. How to be a semantic and existence deflationist, without being a metaontological deflationist

Lorenzo Azzano*
University of Santiago de Compostela

Thomasson (2014) submits that semantic deflationism entails existence deflationism, which in turn entails metaontological deflationism; Marschall and Schindler (2021) argue against this claim. Rather than chiming in on this debate directly, I would like to offer a counterexample to Thomasson's position: I will positively illustrate a position which is not metaontologically deflationary but is surely open to deflationism concerning truth and existence –one which in fact naturally suggests a deflationist reading of such notions. This very limited exploration of logical space aims to show that there are many points which can be conceded to deflationist positions without compromising the central claim that at least some debates in contemporary ontology and metaphysics constitute substantive investigation of an objective and mind-independent reality. In light of this claim, I do not expect to convert any metaontological deflationist: my goal is merely to help them delineate more clearly their adversary, and, by contrast, their own position.

The starting point for this position is the notion of *metaphysical structure*, understood not in a Russellian fact-matching way, but through Sider's (2011) "non-ontic" extension of Lewisian naturalness. Through the use of a syntactically polymorphic joint-carving operator, we may define a metaphysically perspicuous sentence as one formulated in joint-carving terms; viz. a sentence which is not only true, but metaphysically perspicuous in the sense that it reveals, rather than obfuscates, metaphysical structure. As first suggested in Sider (2014: 416) we can reject the material equivalence between truth and perspicuity in order to preserve the substantiality of metaphysical debates from more ordinary everyday beliefs; under this interpretation, debates about realism and anti-realism concerning e.g. numbers, properties, or moral facts, should not be conducted by evaluating the truth of the relevant existential sentences, but their metaphysical perspicuity. Metaphysical investigation, as a substantial investigation about the make-up of objective reality, is then recast as an investigation concerning metaphysical structure.

The resulting philosophical methodology is loosely Quinean in character, in the sense that we should not read our ontology and metaphysics off of just any sentence, but only

* lorenzo.azzano@usc.es

those of a privileged subset (in this case, the metaphysically perspicuous ones –of course, how to settle which sentences are perspicuous is another issue that this methodology has to address); case in point, this is exactly the kind of methodology that a meta-ontological deflationist would oppose.

For our current purposes it is crucial to notice that both truth and existence deflationism are viable options on this account, and, in a sense, constitute a natural extension of the position. First of all, the separation of truth and perspicuity allows us to reject a fact-matching or correspondence theory of truth, according to which is the metaphysical job of truth to “carve up” a piece of reality, presumably in the form of a fact. Nor should we expect to take metaphysically seriously standard Tarskian truth-conditions, for there is no clear indication that such statements in the semantic meta-language are themselves metaphysically perspicuous –that would be the job of what Sider (2011: 132ff) calls “metaphysical semantics”. On the contrary, given the existence of non-perspicuous or “mere” truths, it would seem that the only job of truth is that of expressing factuality (viz. they merely express what is the case) as opposed to metaphysical structure. Since evidently “it is the case that...” is just as redundant as “it is true that...” the value of truth-statements is that they allow one to inject their content into object-language discourse; and although the account has never been explicitly articulated, this would appear to be in the ballpark of semantic deflationism.

Secondly, because mere truths (or their existential generalizations) are not sufficient to settle matters of ontology, this account affords a loosely Carnapian permissivist attitude towards existence (as in Schaffer2009). Although metaphysics, as in investigation on structure, is a substantial and non-trivial endeavor, the existence of lightweight readings of the existential quantifier means that matters of existence are indeed quite trivial on this account, and it is very likely that no non-trivial criterion for existence can be found. A little more grandiosely, on this account is not the job of the metaphysician to find out what exists according to a substantial criterion, nor it is their job to discern true sentences from false ones. Existents and truths are too easy to find. On the contrary, the metaphysician’s investigation on objective reality –*qua* investigation of structure– crucially revolves around the selection of a subclass of true sentences, viz. the metaphysically perspicuous ones.

References

- [1] Marschall, B. & Schindler, T. (2021). Semantic Deflationism and Meta-Ontological Deflationism. *Philosophical Quarterly* 71(1): 99-119. Schaffer, Jonathan (2009). On what grounds what. In David Manley, David J. Chalmers & Ryan Wasserman (eds.), *Metametaphysics: New Essays on the Foundations of Ontology*. Oxford University Press: 347-383.

- [2] Sider, T. (2011). *Writing the Book of the World*. Oxford University Press.
- [3] Sider, T. (2014). Outscoping and Discourse Threat. *Inquiry: An Interdisciplinary Journal of Philosophy* 57(4): 413-426. Thomasson, A. (2014). "Deflationism in Semantics and Metaphysics". In A. Burgess and B. Sherman (eds) *Metasemantics: New Essays on the Foundations of Meaning*, OUP

Rayo's view (2013) from a Deflationist perspective

Concha Martínez Vidal*
University of Santiago de Compostela

Thomasson (2014) characterizes and discusses the relations between different sorts of deflationism. She discusses the relationship between deflationism about truth and existence to claim that the former –deflationism about truth– implies deflationism about existence and that deflationism about existence implies metaontological deflationism (the view that contends that “something is wrong with a great many standard debates in ontology” (Idem 186)). She characterizes deflationism about *X* as any view about *X* that “[...] den[ies] that the relevant concept (of truth, reference, or existence) is even attempting to refer to a substantive property the nature of which we can investigate and hope to discover. As a result, deflationary theories renounce the search for reductive generalizations [...]. Instead, each holds that the relevant concept may be grasped by grasping certain trivial platitudes governing the concept.” (Thomasson 2014, 185)

The purpose of this contribution is to analyze Rayo's 2013 proposal to see to what extent and concerning which notions he can be considered a deflationist. We expect the assessment i) to help understand Rayo's view and its implications; ii) to clarify the relations between deflationism about different concepts. We will pay special attention to his conceptions of reference and truth and their implications concerning deflationism about different notions. Understanding in-depth his commitments in relation to these notions and how those aspects of his view that are deflationist relate to them promises to be especially illuminating.

Rayo himself does not use the term ‘deflationism’ to characterize any aspect of his view. Nevertheless, several features of the view are deflationist. Linnebo (2012) includes Rayo's proposal among the ones he identifies as metaontologically minimalist. Linnebo (2012, 139) characterizes ‘Metaontological Minimalism’ as the view that there are thin objects, objects such that “very little is needed for their existence,” objects that make no demand on the world. Take, for instance, abstract objects like numbers. Rayo thinks that numbers make no demand on the world; this shows in his advocating for (Numbers): For the number of *F*s to be *n* just is for there to be exactly *n* *F*s. (Rayo 2013, 26) Supporting (Numbers) conveys that both the notion of existence (internal to a given logical space)

* mconcepcion.martinez@usc.es

and the notion of an object allow for existing objects in a non-substantive sense of existence and of objecthood; hence Rayo counts as a deflationist about both notions.

Now, according to Marshall and Schlinder’s characterization of metaontological deflationism (not exactly Thomasson’s version above) –“existence has no deep nature and the standard neo-Quinean approach to ontology is misguided” (2021, 99)–, Rayo is a metaontological minimalist too: he agrees existence has no deep nature, he rejects Quine’s neo-Quinean approach since he takes Quine to argue that “the greater a sentence’s ontological commitments, the greater the demands its truth imposes on the world” (Moruzzi & Sereni 2015), and he advocates for a free logic.

Sud and Manley (2021, 101) put forward a positive characterization of a metaontological deflationist view they name ‘Quantifier Variance;’ by this term they refer to the view in general, not just a particular form such as Hirsch’s. Two of those features, namely pluralism (different answers to an existential debate can be true) and egalitarianism (none of the answers is metaphysically distinguished), seem to apply to views that are deflationist about existence though not committed to “Quantifier Variance” such as Rayo’s. Rayo advocates for a version of the Fregean notion of recarving content in terms of his “‘just is’-statements”; (Numbers) above is one of them. ‘Just is’-statements define one’s logical space, characterize the space of logical possibility and come to be accepted or rejected for pragmatic reasons. He is a pluralist about answers to existential debates and about the choice of logical space. He can also be seen as supporting egalitarianism since he denies the objectivity of one’s choice of logical space.

Rayo presents his view as a rejection of “metaphysicalism”, the view according to which:

- (i) “for an atomic sentence to be true, there needs to be a certain kind of correspondence between the logical form of the sentence and the ‘metaphysical structure’ of reality”; and
- (ii) there is a privileged way of carving reality at its joints. His rejection of Metaphysicalism seems motivated by his way of conceiving ‘recarving of content’ in terms of ‘just is’-statements; the latter somehow motivate his “Compositionalism.” His Compositionalism intends to go beyond Frege’s dictum (be an object is to be the referent of a singular term) to explicate what it takes for a singular term to refer: “Let c be a singular term of L and let $[\dots]$ be an interpretation of L . The following conditions are jointly sufficient for c to refer, as interpreted by $[\dots]$: 1) $[\dots]$ preserves logical entailments; 2) $[\exists x(x = c)]$ is a true (coarse-grained) proposition.” (Rayo 2020, 2).

He acknowledges that Compositionalism “does not commit one to the acceptance of any interesting ‘just is’-statements, [though] it does *eliminate an obstacle* for the acceptance of ‘just is’-statements” (Idem, 18), and is compatible with moderate metaphysicalism.

The reason seems to be that Compositionism does not imply ‘egalitarianism’ given that it is compatible with moderate metaphysicalism (Rayo 2013, 11-13).

Discussion of the relations between Compositionism and moderate metaphysicalism brings in the issues of truth and reference, and then, at a different level, the resulting relations between his conception of these notions and the corresponding deflationist views. The notion of object is relative to a language, and his second requirement for Compositionism establishes that $[\exists x(x=c)]$ is to be a true (coarse-grained) proposition, so, the underlying notion of truth cannot imply structural correspondence, but does he need a deflationist notion of truth? What does it exactly mean for $[\exists x(x=c)]$ to be a true (coarse-grained) proposition? Is it a pragmatic question to be decided before subscribing to a given ‘just is’-statement? How do things work internally? What does this say about the links between metaontological deflationism and his form of deflationism about reference and truth?

Fictionalism about Fictional Characters (Reloaded)

Manuel García-Carpintero*
University of Barcelona

In previous work (REDACTED) I offered a fictionalist account of apparent reference to fictional characters in fictional discourse. Here I want to set it in contrast with deflationary realist accounts like Thomasson's (2014). I will present in my own terms how I think best to understand the deflationary realist accounts I target, my alternative fictionalist proposal, and finally I'll outline the reasons to prefer it that I'll develop in my presentation.

Realists about fictional characters argue that a proper account of claims in (1) below requires interpreting names like 'Gregor Samsa' in them as referring to fictional entities:

(1) Gregor Samsa is a fictional character.

Artifactualists like Salmon (1998), Thomasson (1999, 2003) or Schiffer (2003) think of them as having an ontological status analogous to that of the fictional works in which they occur (Thomasson 1999, 143; 2003, 220; Salmon 1998, 78-9). In contrast, my brand of fictionalism about fictional characters claims that we don't need to take referential expressions in such discourses to really refer to them. But I don't have any ontological scruples about *ficta* understood along the lines of the artifactualist proposal. In fact, to assume them is helpful to present my own view. Let me elaborate on how I understand the realist proposal.

Fictional works result in my view (REDACTED) from the communicative acts of fiction-makers; they are social constructs, abstract created artefacts with norm-regulated functions. They have a complex structure, grounded on that of the vehicles that express them; they are in part composed of singular representations (REDACTED). It is these singular representations what I'll take fictional characters to be on the artifactualist account I'll use as a foil: terms like 'Gregor Samsa' have as semantic value a singular representation associated with that very name, which is a constituent of Kafka's *Metamorphosis*.

I'll provide a few more details on how I suggest to think of the semantic values of *prima facie* empty names on the just outlined realist proposals. I assume a view standard in current semantics (REDACTED). On this view, referential expressions like indexicals and proper names carry presuppositions of acquaintance, or familiarity. This is to be

* m.garciacarpintero@ub.edu

cashed out by assuming that contexts include discourse referents, which we may think of as shareable singular representations that may well not pick out anything. For proper names, the relevant discourse referents are crucially defined by naming practices (distinct ones for the ‘David’ that picks out Lewis and the one that picks out Hume, cf. REDACTED); typically existing ones, but in some cases practices created with the very discourse including the name. For indexicals, they might be constituted by perceptual information, or by information available in previous discourse to which the expression is anaphorically linked.

The singular representations that fictional realism takes to be the semantic value of referential expressions like ‘Gregor Samsa’ are thus to be individuated by such discourse referents. There is a tradition that associates some descriptions with entities of the kind we are positing, roles (see Rothschild 2007 and Glavaničová 2017) like the president of the USA or the mayor. Roles can be understood as Carnapian individual concepts picking out their occupiers relative to worlds, to the extent that we think of them as merely partial functions. If we model the fictional world by means of standard possible worlds, the role that we are taking as the semantic value of ‘Gregor Samsa’ will pick out different individuals in different such worlds.

As said, I agree that we are committed to representational entities of the kind just outlined, but I reject that ‘Gregor Samsa’ in (1) refers to one of them. On the view of such discourse I hold (a version of Yablo’s (2001) figuralist brand of fictionalism), the semantic referential apparatus (de jure directly referential expressions such as names and indexicals, quantifiers generalizing over the positions they occupy, expressions for identity) is used metaphorically in the likes of ‘Gregor Samsa is a fictional character’, deploying the figure of speech called hypostatization (REDACTED). It is a rather dead, conventionalized kind of metaphor, so, in contrast to pretense-theoretic fictionalist proposals (Walton 1990, Everett 2013), on this view utterances in metatextual discourse are straightforward assertions with truth-conditions.

This might suggest that the view is after all realist, committed to referents of some sort for singular terms in metatextual discourse, but it is not. In support of this, I follow Yablo’s (2014) recent development of his views, on which the truth of metatextual sentences including fictional names and their generalizations do not really commit us to the existence of fictional characters. For this is merely pretend-presupposed and, when we look at what they are really about (the truth-makers for the claims we make with them), we do not find referents that they pick out.

What we find in all is a fiction like *Metamorphosis*, referentially deploying ‘Gregor Samsa’. This account thus has some similarities with artifactualist views; it relies on similar truth-makers for the likes of (1). We end up interpreting them as making genuine

assertions, whose truth is grounded on the pretenses thereof in the relevant fictional work. But it doesn't need to assign a referent to fictional singular terms like 'Gregor Samsa'.

Why not endorse the deflationary realist proposal? This is not needed on the view just outlined, and there are good reasons against it. The ones I will develop are metaphilosophical in nature. If we take it seriously that fictional names refer, we have to find reasons to choose among the different candidates that realists offer for the role of ficta – Platonic abstracta, Meinongian nonexistents, or what have you, in addition to artifacts; but there is no rational way of taking a stance on such matters, I'll argue. For all we can tell, the relevant empirical and theoretical considerations we can deploy do not select just one of them, even if they do allow us to dispose of some. The fictionalist attitude authorizes us to ignore the issue in good faith.

Indexicals, the self and time

María Ponte Azcárate*
University of the Basque Country

Philosophical discussions involving indexicals often involve ontological theses, either as assumptions taken as fundamental for a correct semantic account or, quite often, as alleged consequences of them. One paradigmatic case is Arthur Prior's discussion of the temporal indexical "now" in his 1959 paper "Thank Goodness that's over" (TG). Another is John Perry's analysis of the first-person pronoun "I", in his 1977 paper "The problem of the essential indexical". Prior's paper is normally interpreted as defending the existence of the A-series of time, according to which past, present and future are real properties of reality. Perry's paper has been used to defend the essentiality and existence of the so-called *de se* thoughts, which involve the self-attribution of certain properties, thus requiring a rich ontology regarding properties and the self. Neither paper, I claim, should be read as involving ontological theses. Both papers discuss the importance of indexicals in language, thought and action. But the explanation of these indexicals and the role they play does not, or should not, require postulating particular ontological objects, properties, events or relations. Alleged ontological issues derived from the study of indexicals are, I argue, nonsubstantive, in the sense that nothing substantial is at stake, beyond semantic, pragmatic and epistemological matters. In discussions of the role of time in language and thought, it is often unclear how the various aspects of what is being said are related to each other. Ontological considerations about the nature of time are raised, together with epistemological (and phenomenological) aspects of how we experience the flow and the direction of time. Quite often, semantic (and pragmatic) insights into our talk about time are also mentioned, without clear connections and distinctions between the arguments at each level. Prior's TG is a clear case at hand: ontological, epistemic, and linguistic considerations are entangled in a way that creates the illusion of an ontological argument about the nature of time. Prior's arguments are normally interpreted as ontological: a defence of the reality of McTaggart's A-series. To understand what the A-series of time is, and its difference with the B-series, consider the following questions: Is there an objective difference between past, present and future? Are they pieces of reality's furniture or just parts of human's cognitive and linguistic apparatus? The A-theory of time claims that there is an objective difference and that past, present and future are "real" properties of reality. The B-theory denies it. According to the standard interpretation, Prior's main

* maria.deponte@ehu.es

claim in TG is that B theories lack the tools to warrant thoughts about past events and the emotions they elicit –regret, relief, etc.–and the essential differences with thoughts about future events and the emotions they elicit–anxiety, hope, etc. Accordingly, a view that denies the existence of A-properties offers no grounds for tensed thoughts and tensed emotions. But there is no clear indication in TG to conclude that Prior is discussing ontological issues, or postulating the existence of A-properties or A-series of time. There is no clear indication, either in TG or elsewhere, that a defence of A-properties² is necessary to explain tensed thoughts, the temporal indexicals we normally use to express them, and the different cognitive reactions temporal events elicit in us. I first question this interpretation of TG, based on a careful reading of the text and a discussion Prior held with Cohen about it right after its publication. I claim that, in TG, Prior is discussing ways to understand and talk about time in ordinary language and, more to the point, that he is dealing with issues of cognitive significance involved in our uses of temporal indexicals such as “now”. In this sense, Prior is clearly anticipating the discussion about indexicals present in Perry (1977). I then argue that a Perry-inspired framework offers a more accurate and ontologically simpler reading of TG. In his 1977 paper, Perry discusses the essential role indexicals play in thought and action, focusing in the indexical “I”, but also discussing “now”. Recently, Perry’s views have been challenged (Millikan, 1990, Cappelen and Dever, 2013, Magidor, 2015), claiming that the alleged essentiality of indexicals and the related notion of *de se* attitudes are just a “myth”. I focus on temporal indexicals and I defend that they, and what Perry calls “self-locating beliefs”, play an important role in our accounts of cognition and action. In doing so, neither I nor Perry (as far as I know) commit to *de se* or *de nunc* attitudes, understood as involving the self-attribution of properties or self-attached knowledge. *De se* and *de nunc* attitudes might perfectly well be a myth, if they are understood as something other than self-locating beliefs, but the role indexicals play in thought and in our explanations of action is not. Indexicals are essential, according to Perry, only in the sense that they point at 3 different *hows*, different ways to believe an agent might have. Shortly put, the difference between the beliefs “I am writing” and “María is writing”, assuming I am the agent in both cases, relies not in what I believe, but in how I do so. Similarly, the difference between exclaiming, after a painful experience, “Thank goodness it is over (now)!” and “Thank goodness it is over by April 7th 2021!” relies in how, and not in what we are thankful for. With regard to the ontology of time, I conclude that there is indeed a difference between A-series representations and B-series representations, but that they represent the “same things”. McTaggart’s A-series and B-series are the same series, described and thought about in different ways. Thus I claim that the difference between the A representation and the B representation is not what is represented as happening, but how it is represented. In other words, the difference is not ontological, in any substantial way.

References

- [1] Cappelen, Herman and Josh Dever. *The Myth of the Essential Indexical* Oxford, Oxford University Press, 2013.
- [2] Magidor, Ofra (2015). *The Myth of the De Se*. *Philosophical Perspectives* 29 (1): 249- 283.
- [3] Millikan, Ruth (1990). *The Myth of the Essential Indexical*. *Nous*, 24(5), 723-734.
Perry, John (1977). "Frege on Demonstratives". *Philosophical Review*, 86, no. 4: 474-97.
Prior, Arthur (1959) "Thank Goodness That is Over." *Philosophy* 34:12-17

SYMPOSIUM:
PHILOSOPHIES OF SCIENTIFIC
EXPERIMENTATION: OLD AND
NEW CHALLENGES

Abstract

Ever since the 1980s philosophical interest over scientific experimentation have been rife. An important number of studies have explored different settings, technologies and practices as related to experimental performance and validation. The interest over laboratories, experimental systems, and the entities there created and manipulated (whether elementary particles, genes, model organisms or subjects in social experiments) have open a window on new problems and questions over the nature of scientific knowledge. This symposium aims to explore and discuss some of these problems and questions, especially in relation to Biochemistry, Evolutionary Biology and Social Sciences. The first speaker will examine the concept and process of “purification” of biochemical entities within the context of *in vitro/in vivo*’s distinction in experimental biology. The second speaker will explore the connection between modeling and experimentation in evolutionary biology arguing for the need to distinguish between system of phenomena and target systems. The third speaker will focus on the concept of “reactivity” in the context of social sciences (Economics). In particular, she will examine different sorts of “reactivities” – whether positive or negative – when experimenters work with human beings. The last speaker will revisit the epistemic tradition of the “maker’s theory of knowledge” and show how that might illuminate further the nature of experimental activities.

Justification

The Symposium aims to discuss seminal philosophical problems as related to scientific experimentation. Over the last few decades, the philosophy of experiment has become a central and vibrant sub-field within history and philosophy of science. The participant involved in the symposium are all experienced scholars with many years dedicated to study the nature of scientific experimentation. Accordingly, we do think that the Symposium would enrich further the variety of philosophical debates and talks that will be presented in in the X Congress of the SLMFCE.

Purificación de entidades y obtención de factores causales: aspectos epistemológicos de la práctica experimental en bioquímica

Gabriel Vallejos*
Universidad de Chile

El estudio de entidades biológicas purificadas es una de las prácticas más extendidas en la bioquímica, pero es también el origen de su más importante problema epistemológico: el problema *in vitro/in vivo*, que consiste en justificar cómo es posible obtener conocimiento acerca de la naturaleza de los sistemas biológicos estudiando sus partes por separado y en contextos distintos a su medio natural.

Salvo algunas excepciones [1-6] y pese a ser ubicuo en la práctica científica, este problema no ha recibido considerable atención en la literatura filosófica. Es frecuente encontrar que los tratamientos se centren en aspectos representacionales y que generalmente concluyan en forma pesimista ante una posible solución. Por otro lado, se suele tomar como punto de partida a las entidades ya purificadas, sin prestar atención al proceso de purificación y a su instanciación material en el laboratorio. En este trabajo, utilizando como ejemplo la purificación de proteínas, se mostrará que considerar los procesos de purificación de entidades y de construcción de sistemas experimentales posibilita vislumbrar soluciones al problema *in vitro/in vivo*. Se argumentará que en estos procesos se aíslan factores causales robustos asociados a propiedades de las entidades a purificar, los que luego pueden ser utilizados para la construcción de teorías biológicas y para intervenir causalmente en sistemas a los que estas entidades pertenezcan.

Los sistemas experimentales (SEs) son la unidad mínima de investigación en biología experimental. Consisten en sistemas artificiales de manipulación y control causal, contruidos y estandarizados con el fin de obtener un fenómeno estable y producir resultados confiables y replicables. Cada SE posee un mecanismo interno que subyace a los procesos mediante los que se ejerce el control del sistema, que está dado por las interacciones entre la(s) entidad(es) biológica(s) y los objetos técnicos que lo conforman.

Todos los SEs forman parte de una compleja red material de experimentación compuesta por muchos SEs en mutua dependencia. Por ejemplo, muchos ingredientes para la construcción de un SE son producidos mediante otros SEs en lo que se conoce como “experimentación preparativa” en la literatura filosófica [7-9]. Esto incluye los procesos de purificación (PP) de entidades biológicas, compuestos por varios SEs sucesivos. En

* gvallejos@ug.uchile.cl

nuestro ejemplo, se comienza con la lisis celular de un cultivo bacteriano para producir un extracto crudo, el que luego es sometido a diversos SEs tales como cromatografías, precipitación selectiva, diálisis, concentración, etc. hasta obtener una muestra final de proteína pura.

En cada una de las etapas de un PP se produce el aislamiento de factores causales (FCs) específicos asociados a propiedades de la entidad que se está purificando, tales como su tamaño, forma, carga eléctrica, solubilidad, estabilidad térmica, capacidad de interacción con otras moléculas, actividad enzimática, etc. Estos FC, que forman parte del mecanismo interno del SE, son los que garantizan su control causal mediante la manipulación robusta y exitosa de la entidad a purificar. Así mismo, estos FC también garantizan el control exitoso de una gran variedad de SEs donde la entidad purificada (o a purificar) está presente. Es por eso que dudar de la existencia de estos FC conllevaría el costo de dudar de todos los procesos de manipulación y control que posibilitan.

Teniendo todo esto en cuenta, es natural inferir que los FC utilizados para manipular robustamente una entidad también juegan un rol en los sistemas biológicos de los cuales esa entidad forma parte en la naturaleza, siendo así incluidos en teorías y modelos acerca de mecanismos biológicos. De esta forma, considerando el proceso de purificación de entidades, es posible proponer soluciones para el problema *in vitro/in vivo* en la biología experimental.

Bibliografía

- [4] Rheinberger H-J. (1997) *Toward a History of Epistemic Things*, Stanford University Press.
- [5] Strand R, Fjelland R & Flatmark T. (1996) "In vivo interpretation of in vitro effect studies with a detailed analysis of the method of in vitro transcription in isolated cell nuclei". *Acta Biotheoretica* 44: 1-21.
- [6] Strand R. (1999) "Towards a useful philosophy of biochemistry: Sketches and examples". *Foundations of Chemistry* 1 (3): 269-292.
- [7] Jacob C. (2002) "Philosophy and biochemistry: Research at the interface between chemistry and biology". *Foundations of Chemistry* 4 (2): 97-125.
- [8] Ibarra A & Mormann T. (2006) "Scientific Theories as Intervening Representations" *Theoria* 21, N° 55: 21-38.
- [9] García P. (2015) "Computer simulations and experiments: in vivo–in vitro conditions in biochemistry". *Foundations of Chemistry* 17 (1):49-65.

- [10] Weber M. (2005) *Philosophy of Experimental Biology*, Cambridge University Press
- [11] Waters CK (2008) “How Practical Know-How Contextualizes Theoretical Knowledge: Exporting Causal Knowledge From Laboratory to Nature”. *Philosophy of Science* 75 (5):707-719.
- [12] Esposito M. & Vallejos G. (2020) “Performative Epistemology and the Philosophy of Experimental Biology: A Synoptic Overview”. en Baravalle L. & Zaterka L. (eds.) *Life and Evolution - Latin American Essays on the History and Philosophy of Biology*. Springer.

Modeling, explanation, and experimentation in evolutionary biology: the distinction between phenomena and target system

Rodrigo López Orellana*
Universidad de Valparaíso

Our talk aims to present a perspective on modeling and experimentation in evolutionary biology based on the distinction between phenomena and target systems. Our proposal follows, on onehand, the idea that the target of any model is always the outcome of a complex process of construction that should be differentiated from phenomena (Cassini, 2018); on the other hand, our proposal is framed within a pragmatic and dynamic approach to modeling in biology focused on the interactive role of target systems (Redmond, 2021). This interactive role, in our view, corresponds to the inferential process triggered in the practice of scientific modeling (i.e., surrogate reasoning), and which is captured by the definition of scientific explanation. In effect, we identify the target systems as one of the three axes of the modeling process, together with models and phenomena. As Godfrey-Smith (2006, 726) stated, modeling should be understood as a strategy of indirect representation of the world. Such a strategy allows us to understand complex actual systems through the understanding of a hypothetical system—a model—but similar to it in some relevant aspects. However, we believe that there is an intermediate step of modeling construction in which the target system is produced as another side of the same product. The representation of a real phenomenon is also indirect in this sense, the model indirectly points to the real system but directly to a target system. Thus, a target system is also a product—in the same way that the model is produced and created—, i.e., it is an interactive component of the inferential dynamics that characterizes the scientific practice with explanatory models. Hence, our inferential approach distinguishes, both epistemologically and ontologically, between the target systems and the phenomena.

Furthermore, we argue that—especially in the case of biology— experiment and the artificial production of phenomena play an important role in the dynamics of modeling. Following Hacking (1983, 263-264), we believe that model and experiment are closely related. The function of experiments is not reduced to the confirmation or refutation of theories. Indeed, Hacking considers experimentation as a dynamic practice—as action— based on the intervention—manipulation— and transformation of phenomena. This practice has a certain autonomy from the theory(s), as does the practice with models

* rodrigo.lopez@usal.es

(Morrison y Morgan, 1999b, 10), and from it, the distinction between phenomena and target systems can be better understood.

To fulfill our objective, we will apply our approach to an evolutionary developmental biology (evo- devo) model: the Polypterus (Senegal bichir), an organism-model used to explain the role of developmental plasticity in the evolutionary origin of tetrapods in the late Devonian (Standen, Du, and Larsson, 2014). This model is highly exemplary for accounting for the inferential process of modeling complex biological systems. As is well known, evo-devo is an area of biology that remains at the center of the current academic debate concerning the epistemological problems it presents, centered on its lack of explanatory and descriptive mechanisms of genetic assimilation that would guarantee the validity of its inferences in macroevolutionary terms; evo-devo is claimed to offer only explanations based on local experimental conclusions (which would be very limited). We argue here, that if we distinguish between the system of phenomena and the target system, we can establish some criteria by which this model would be able to provide us with a good explanation of the phenomena and thus gain cognitive success from the experiment with this organism (*Polypterus senegalus*).

References

- [1] Cassini, A. (2018). Models without a target. *ArtefaCToS*, 7(2), 185-209. <https://doi.org/10.14201/art201872185209>
- [2] Godfrey-Smith, P. (2006). The strategy of model-based science. *Biology and Philosophy*, 21, 725–740. <https://doi.org/10.1007/s10539-006-9054-6>
- [3] Redmond, J. (2021-Online first). A free dialogical logic for surrogate reasoning: generation of hypothesis without ontological commitments. *THEORIA. An International Journal for Theory, History and Foundations of Science*. <https://doi.org/10.1387/theoria.21902>
- [4] Hacking, I. (1983). *Representing and Intervening. Introductory Topics in the Philosophy of Natural Science*. Cambridge: Cambridge University Press.
- [5] Morrison, M., & Morgan, M. S. (1999b). Models as mediating instruments. In M. Morrison & M. S. Morgan (Eds.), *Models as Mediators. Perspectives on Natural and Social Science* (pp. 10-37). Cambridge: Cambridge University Press.
- [6] Standen, E. M., Du, T. Y., & Larsson, H. C. E. (2014). Developmental plasticity and the origin of tetrapods. *Nature*, 513(7516), 54–58. <https://doi.org/10.1038/nature13708>

Reactividad y Validez experimental

María Jiménez-Buedo*
UNED

El auge de la experimentación científica social de los últimos años ha sido impulsado en gran parte por el éxito de la economía experimental y del comportamiento. Por tanto, es natural que las discusiones metodológicas en torno a las nuevas prácticas experimentales en las ciencias sociales hayan sido a menudo moldeadas por los debates que estaban teniendo lugar entre los economistas experimentales en ejercicio.

Esto puede explicar, al menos en parte, por qué, durante años, la reactividad, o el fenómeno que ocurre cuando los individuos alteran su comportamiento debido a su conciencia de ser estudiados, pareció preocupar poco o nada a los economistas experimentales. Con su clara postura metodológicamoldeada por la lista de preceptos de Vernon Smith (Smith 1982), los economistas pudieron evitar las preocupaciones asociadas con la reactividad de los sujetos mediante un control tenaz sobre los incentivos propuestos por el escenario experimental.

Más recientemente, a medida que los economistas experimentales avanzaban gradualmente en su estudio hacia temas en los que los incentivos económicos se mezclaban con consideraciones normativas (como en el estudio del altruismo, el castigo o las normas sociales), ha surgido un interés correspondiente en el problema de la reactividad (Zizzo 2008, pág. Bardsley 2007, Levitt y List 2011).

Una vez interesados en el fenómeno de la reactividad, los economistas, en su mayor parte, han buscado en los cajones conceptuales de sus colegas psicólogos sociales y han recogido una variedad de términos que son de uso común en psicología. Estos términos, y sus usos, suelen, no obstante compartir problemas de ambigüedad conceptual: Los efectos de Hawthorne, los efectos de demanda de la experimentación, o los efectos placebo son todos ellos términos que a menudo se usan indistintamente entre sí, y, lo que es más importante, se usan tanto para referirse al fenómeno de reactividad como a los efectos de sesgo potenciales o reales que tiene este fenómeno. En otras palabras, estos términos a menudo se usan de manera indistinguible tanto para referirse al fenómeno por el cual los sujetos alteran su comportamiento como resultado de ser observados como al hecho de que esta alteración tiene el potencial de sesgar los resultados de los experimentos.

De esta forma, a referencia a la reactividad experimental puede sugerir cosas muy diferentes, en términos de cuán amenazadores pueden ser para la validez experimental: desde la visión metodológica de los psicólogos sociales, para los que la reactividad es

* mjbuedo@fsof.uned.es

algo para la que los experimentalistas tienen procedimientos rudimentarios que la previenen o detectan, hasta las visiones ligadas a algunas formas de anti-naturalismo, en la que la reactividad ha representado, para otros, un obstáculo definitivo a la mera posibilidad de investigar lo social de manera experimental (Harré y Secord, 1971).

En esta presentación utilizamos un marco intervencionista (Woodward 2003) para analizar los desafíos que la reactividad puede plantear a la inferencia causal en experimentos con humanos. Este marco nos permite distinguir entre formas de reactividad malignas y benignas, en términos de sus efectos sobre la validez de las inferencias causales extraídas de los resultados experimentales. Además, argumentamos que las formas malignas de reactividad no hacen que los datos de un experimento dado carezcan de interés, pero sí hacen que sean causalmente ininterpretables.

Bibliografía

- [1] Bardsley, N. (2008). Dictator game giving: altruism or artefact?. *Experimental Economics*, 11(2), 122-133.
- [2] Harré, R., & Secord, P. F. (1972). *The explanation of social behaviour*.
- [3] Levitt, S. D., & List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world?. *Journal of Economic perspectives*, 21(2), 153-174.
- [4] Woodward, J. (2003). *Making things happen: A causal theory of explanation*. Oxford: Oxford University Press.
- [5] Zizzo, D. J. (2010). Experimenter demand effects in economic experiments. *Experimental Economics*, 13(1), 75-98.

La filosofía del experimento y la teoría poiética del conocimiento: ¿Una relación constitutiva?

Maurizio Esposito*
University of Lisbon

Es bien sabido que Ian Hacking criticó la actitud de subsumir el experimento a la teorización científica. En la década de los 80s, el filósofo canadiense reclamó que los filósofos de la ciencia habían dedicado atención exclusiva a las teorías científicas olvidándose la actividad experimental y de las tecnologías relacionadas. Hacking argumentó que la actividad experimental tenía una vida propia que no podía reducirse a una mera estrategia de comprobación teórica (Hacking 1982). Si bien el giro experimental de los años 80 enriqueció las discusiones filosóficas sobre la naturaleza del conocimiento científico - permitiendo reconsiderar la dicotomía realismo/antirrealismo y estimulando nuevas generaciones en dedicar mayor atención a lo que ocurría en los laboratorios - sus consecuencias han sido relativamente modestas. En primer lugar, contrariamente a la famosa reclamación de Hacking, los filósofos ya se habían interesado en el “experimento” hace décadas, si bien no lo habían siempre hecho en idioma inglés (Simons, 2021). En segundo lugar, la filosofía de experimento de Hacking no permitía superar aquellas distinciones filosóficas que supuestamente debía superar. Dicotomías como realismo/antirrealismo, concreto/abstracto, teoría/práctica fueron sin dudas reformuladas, pero no podían ser trascendidas sin una reconsideración radical de la epistemología subyacente que hacía posible estas mismas distinciones.

Unos de los límites principales en las discusiones que conformaron el así llamado giro experimentalista fue la falta de una seria consideración de la que en el mundo anglosajón es conocida como “the maker’s theory of knowledge” y que aquí tradujo con teoría poiética del conocimiento. La historia de esta tradición es larga. Encontramos sus primeros defensores en el contexto de la filosofía helenística, medieval y renacentista (Mondolfo 1971). En el periodo moderno Francis Bacon y Giambattista Vico ofrecieron su versión más sofisticada (Perez-Ramos 1988). En el 1710, en contradicción con una tradición histórica milenaria (platónica y aristotélica) y, al mismo tiempo, cuestionando el racionalismo cartesiano, Vico propuso que el conocer no se relacionaba ni con creencias, ni con representaciones y tampoco con ideas claras y distintas; el conocer debía vincularse con la acción, y entonces producción y operación. La operación epistémica podía referirse a entidades abstractas y materiales, fueran estos números, figuras geométricas, o entidades físicas (artefactos). Además, la filosofía del conocimiento vichiana no identificaba la ciencia con

* mauriespo@gmail.com

la razón, sino con la facultad del ingenio (*Ingenium*). El *ingenium* vichiano consistía en aquella habilidad *práctica* de identificar las relaciones causales entre las varias entidades del mundo natural o artificial. En términos muy pragmáticos, Vico concibió la noción de “causa” en términos de “operación” (*Negotium*), y con “operación” Vico se refería a todas aquellas actividades humanas que producen efectos (Vico 2008). En otras palabras, las relaciones lógico/causales que atribuimos a los fenómenos concretos sólo emergen a partir de una *operación* cognitiva o concreta del investigador, el cual genera, intencional o inintencionalmente, efectos específicos con su propia actividad. Conocemos un teorema porque generamos sus axiomas y conocemos un sistema experimental porque lo constituimos y utilizamos. No sorprendentemente, Vico consideraba la física galileana como el prototipo de conocimiento poiético. El “experimento” galileano no tenía una función instrumental para testear teorías, sino que era una “operación” que reproducía un fenómeno supuestamente natural modelándolo en un espacio artificial.

La teoría poiética del conocimiento ha sido recientemente reconsiderada por Luciano Floridi (Pezzano 2019), el cual ha diferenciado entre dos tipos de enfoques epistémicos: un enfoque basado en el usuario de información (*user’s knowledge approach*) y un enfoque poiético (*maker’s knowledge approach*) (Floridi, 2011). El primer enfoque es el más antiguo y está orientado por la idea que conocemos como usuarios pasivos de información derivada. Por tanto, nuestra actividad se reduce a una evaluación y justificación de dicha información. En el segundo enfoque, por lo contrario, conocemos como agentes activos que producen información (y artefactos) y esto caracteriza la perspectiva poiética ya mencionada.

Sobre la base de este marco conceptual, en la ponencia pretendo defender la oportunidad de vincular la filosofía del experimento con la teoría poiética de conocimiento considerando los siguientes puntos:

- a) La actividad experimental, como actividad poiética, supone una epistemología constructivista (vichiana o floridiana). Los límites y alcances del conocimiento experimental dependen esencialmente de nuestras actividades.
- b) La diferencia entre abstracción y hacer concreto (*practica*) es solo una distinción de grado. Una teoría o una abstracción es tan real como un objeto físico utilizado para generar efectos en un sistema experimental. Dicho en otras palabras, teorizar es una práctica y no es necesario definir la primera en oposición a la segunda.
- c) La distinción entre realismo y antirrealismo solo subsiste si acepta un enfoque “*user*” y no poiético del conocimiento. En el contexto de una perspectiva poiética, la distinción entre objetivo y subjetivo se transforma en conocimientos “robustos” y “débiles”. Y lo que se considera “robusto” y “débil” no puede ser objeto de formalización general.
- d) Por tanto, la tarea filosófica no es fundamentar la información derivada de la actividad experimental sino explicitar los elementos epistémicos implícitos que proporcionan dicha actividad. La tarea de fundamentar o justificar conocimientos es una

herencia del enfoque “user”; una herencia que ha producido controversias interminables y distintas formas de escepticismo epistémico.

- e) La perspectiva poética del conocimiento es, al mismo tiempo, constructivista y realista(sin que esto sea, necesariamente, una contradicción).

Referencias

- [1] Floridi L 2011, A Defense of Constructionism: Philosophy as a Conceptual Engineering, *Metaphilosophy*, Vo. 42, N. 3.
- [2] Hacking I, 1982, Experimentation and Scientific Realism, *Philosophical Topics*, Vol. 13, No. 1 Mondolfo R, 1971, *Verum-Factum: Desde antes de Vico hasta Marx*, Siglo XXI Editores.
- [3] Pérez-Ramos A, 1988, *Francis Bacon's Idea of Science and the Maker's Knowledge Tradition*. Oxford: Oxford University Press.
- [4] Pezzano G, 2019, Italian Hacking ed eredità: La ragione costruzionista da Giambattista Vico a Luciano Floridi, *Filosofia Italiana*, VOL. XIV, 2, pp. 89-106.
- [5] Simons M, 2021, Were experiments ever neglected? Ian Hacking and the history of the philosophy of experiments, *Philosophical Inquiries*. 9 (1). p.167-188.
- [6] Vico G, 2008, *Metafisica e Metodo*, Bompiani, Milano.

SIMPOSIO:
CUESTIONES DE
METAFÍSICA HOY

El cambio de siglo ha traído consigo la rehabilitación de la Metafísica en todos los ámbitos de la Filosofía. Tras la caída de las posiciones más positivistas, y en un mundo cada vez más interdisciplinar e interconectado, han cobrado fuerza renovada las preguntas por las cuestiones universales más fundamentales, que llegan en forma de metafísica desde disciplinas como la Filosofía de la Lógica y de la Ciencia, el cruce entre Lenguaje y Ontología, o la Filosofía de la Mente y la Fenomenología. Así, el desarrollo de la lógica modal cuantificada ha llevado a proponer diferentes formas posibles de entender la noción de existencia y la modalidad, a debates en torno al realismo/ficcionalismo de mundos posibles (Lewis, Rosen), a nuevas formas de esencialismo (Fine), a repensar la necesidad de la identidad (Kripke), etc. En Filosofía del Lenguaje, el estudio y el cuestionamiento de los nombres propios y los nombres de clases naturales como designadores rígidos, descripciones, convenciones, etc, o el estudio de los discursos de ficción, han agudizado debates que igualmente replantean muy diversas posiciones entre el realismo (Zalta) y el deflacionismo (Thomasson). El desarrollo de la teoría de los actos de habla ha afectado a múltiples preguntas de rango ontológico, abarcando también la pregunta por la fundamentación, emergencia o formas de existencia de los objetos sociales (Clower). Igualmente la Filosofía de la Mente resucita clásicos problemas sobre identidad con una mirada renovada a la Fenomenología actual y recupera la reflexión sobre la experiencia fenomenológica del cuerpo, además de la revitalización de la intuición como posible vía de acceso a la modalidad (BonJour, Chudnoff).

Por supuesto, el objeto de este simposio no podría ser el considerar todas estas cuestiones, ni siquiera desde un punto de vista global. Se trata, sin embargo, de reunir a 5 investigadores que, en distintas disciplinas filosóficas y distintas universidades en España, consideran su trabajo representativo de la Metafísica Contemporánea teniendo especialmente presente las cuestiones mencionadas y cómo estas afectan a las categorías metafísicas fundamentales de Posibilidad, Existencia, Identidad y Objeto. La puesta en común de su trabajo individual se concretiza en 5 ponencias que atraviesan horizontalmente el estudio de estas categorías y cuestiones de la siguiente forma:

- i) La consideración crítica de la posición esencialista de Fine sobre la identidad de objetos materiales.
- ii) La pregunta por la relación entre los conceptos modales aristotélicos de potencia y actualidad con los de posibilidad y necesidad en la metafísica y lógica modal contemporáneas.
- iii) El estudio crítico del uso de la intuición y de la experiencia de los cuerpos para la comprensión de la identidad.
- iv) La interpretación y defensa de la posición de Kripke sobre la existencia de objetos ficticios y sobre los enunciados de existencia negativa.
- v) Análisis crítico de posiciones emergentistas en la explicación de la propiedad social de la “cambiabilidad” del dinero sobre sus propiedades físicas.

El X congreso de la SLMFCE ofrece una excelente oportunidad para crear un foro horizontal de discusión sobre cuestiones Metafísica Contemporánea en España.

Resumen

El objeto de este simposio es poner en comunicación el trabajo de 5 investigadores que, desde distintas disciplinas filosóficas como la Filosofía de la Lógica, del Lenguaje, de la Mente y la Fenomenología, consideran su investigación representativa de la Metafísica Contemporánea, en concreto, en cuestiones que afectan a la modalidad, la identidad y la existencia de los objetos, tanto materiales como sociales. Se presentarán, por tanto, cinco ponencias que abarcan: la contraposición entre la metafísica modal contemporánea y la aristotélica, la referencia a objetos de ficción y los enunciados de existencia negativos, la consideración crítica de la posición de Fine sobre la existencia objetos materiales, el estudio y estatus ontológico del dinero como objeto social, o el papel de la intuición en la experiencia corpórea.

Sobre la no-existencia de las “entidades” de ficción

María José García-Encinas*
Universidad de Granada

En su conferencia “Vacuous Names and Fictional Entities”, Kripke (1973) concluía con tres posibles respuestas a la cuestión *¿Existe Hamlet?*:

(I) Sí, *Hamlet existe* (y existe contingentemente) porque, de acuerdo a la ficción que Shakespeare escribió, Hamlet piensa, además de hacer otro montón de cosas. En la ficción, Hamlet piensa, luego Hamlet existe. Por otra parte, Gonzago, a su vez un carácter ficticio dentro de la ficción, *no* existe. ‘Hamlet’, dentro de la ficción, no es un nombre propio, sino un pretendido nombre propio. Igualmente, que Hamlet habla con Ofelia no es un hecho o una proposición, sino un hecho pretendido o una proposición pretendida. Así, es una verdad pretendida que Hamlet existe, y que existe contingentemente.

(II) Sí, *Hamlet existe* (y existe contingentemente) porque Hamlet es un *objeto abstracto* creado por la imaginación de Shakespeare. Fuera del discurso de ficción, críticos literarios y filósofos hablan de forma natural sobre este objeto abstracto. Y es, claro, un hecho contingente que Shakespeare creara a Hamlet, luego el objeto abstracto que es Hamlet existe contingentemente.

La tercera respuesta es negativa:

(III) No, *Hamlet no existe* (y es imposible que exista) porque Hamlet es una ficción. No hay nadie en el mundo que haya hecho lo que la tragedia cuenta. Incluso si hubiera habido alguien que lo hubiese hecho, no sería Hamlet. Es decir, contra otras posiciones que entienden que las ficciones son mundos posibles (Lewis, 1978), Kripke niega que las entidades de ficción sean entidades posibles de ningún tipo.

Con estas tres ideas en mente, defenderé *primero* que ‘Hamlet’ no es un nombre propio, ni siquiera en el sentido de (II): Las ficciones no son objetos abstractos; no hay, en ningún sentido, entidades de ficción.

* encinas@ugr.es

Una posición habitual mantiene que ‘Hamlet’ en (I) y ‘Hamlet’ en (II) son dos nombres propios diferentes. Una razón principal para esta afirmación (e.g., Thomasson 1999) es que no es posible predicar de objetos propiedades contradictorias. Si Hamlet en la ficción, y Hamlet como objeto abstracto fuesen el mismo objeto resultaría un objeto con propiedades contradictorias o absurdas. Por ejemplo, Hamlet aún existe, pero murió en el siglo xvi. O, puesto que Hamlet quiere a Ofelia habría que admitir que las entidades abstractas pueden amar a sus hijas. Lo mismo, Hamlet fue creado por Shakespeare, pero sus progenitores fueron el Rey Hamlet y la reina Gertrudis. Etc. Estas parecen por tanto buenas razones para mantener que tenemos aquí dos nombres diferentes: ‘Hamlet’ en la ficción, un nombre pretendido, en realidad, y que pretende referir alguien, según (I). Y ‘Hamlet’, nombre propio que refiere al objeto abstracto creado por Shakespeare.

A pasar de esto, quiero defender que ‘Hamlet’ no es un nombre propio tampoco en el sentido de (II): ‘Hamlet’ no refiere a diferentes entidades en diferentes contextos. En particular, no refiere a ninguna entidad abstracta que Shakespeare creara. Más bien, nunca refiere: Es un término cuyo uso es “hacer de nombre propio” o funcionar “como si” refiriere, tal y como se sigue de (I). Y al usar el lenguaje “como si” hablase del mundo, no creamos un nuevo mundo.

La dificultad es que, si Hamlet no existe, ha de explicarse cómo es posible sentir admiración o lástima por Hamlet. Uno admira a Hamlet: su personalidad, sus acciones, su naturaleza ... Es claro que, admirando a Hamlet, uno no admira un objeto abstracto, ni un objeto intensional creado por la mente. Los objetos abstractos no tienen personalidad, ni hacen cosas. Igualmente, cuando uno siente pena por Hamlet no se convierte en ficción y siente entonces, pretendidamente, pena por Hamlet. No es que, de acuerdo con la historia, uno admire a Hamlet. Y tampoco admira uno una ‘pretensión’. Mi propuesta será, más bien, que uno puede admirar a Hamlet porque ‘Hamlet’ fuera del discurso de ficción *del que es directamente sacado* es una exáfora (una forma de metáfora). En ‘Admiro a Hamlet’, ‘Hamlet’ establece una relación entre quien admira y Hamlet en la ficción: el término hace de vínculo con el personaje, aunque Hamlet no está presente (no existe) en el contexto de su admirador. Hamlet puede así ser admirado, como pueden ser temidas las múltiples caras de la Guerra aun cuando la guerra no tiene caras. Uno puede incluso admirar un Hamlet contradictorio, o temer de verdad como Hamlet teme en la ficción.

Una segunda dificultad es que, si no hay ningún objeto que sea Hamlet, no parece que pueda decirse con verdad que Hamlet *no* existe. Así, en *segundo lugar*, y en el seno de una teoría Kripkeana de los nombres propios según la cual ‘Hamlet’ no tiene significado, defenderé que ‘Hamlet no existe’, o ‘Necesariamente, Hamlet no existe’ son verdadero, *pero* (i) su verdad no implica que hay un objeto que no existe –contra Parsons (1980) o Zalta (1983); y (ii) su verdad no implica que la existencia es deflacionaria y relativa a un contexto –contra Thomasson (2015). Al contrario, siguiendo a Kripke (1973), defenderé

que enunciados como ‘Shakespeare existe’ *no* expresan proposiciones: Expresan hechos de existencia, sí, pero carecen de contenido proposicional. Son enunciados puramente formales que expresan formas estructurales. Igualmente, el enunciado ‘Hamlet no existe’ expresa la ausencia de un hecho: no hay hecho, ni siquiera posible, sobre Hamlet. Los enunciados de existencia son juicios predoxásticos intuitivos de razón, al modo de los juicios intuitivos defendidos por Bonjour (1998) o Chudnoff (2013).

References

- [1] Bonjour L., 1998. *In defense of pure reason: A rationalist account of a priori justification* Cambridge University Press.
- [2] Chudnoff, E., 2013. *Intuition* Oxford University Press.
- [3] Kripke, S., 1973. Vacuous names and fictional entities. Reprinted in 2011, *Philosophical troubles: Collected Papers I*, pp. 52-74, Oxford University Press.
- [4] Lewis, D., 1978. Truth in fiction. *American Philosophical Quarterly* 15 (1): 37-46.
Parsons, T., 1980. *Nonexistent objects* Yale University Press.
- [5] Thomasson, A., 1999. *Fiction and metaphysics* Cambridge University Press. Thomasson, A. 2015., Fictional discourse and fictionalisms. In Everett & Brock (eds.) *Fictional objects* pp. 255-74, Oxford University Press.
- [6] Zalta, E., 1983. *Abstract Objects* Reidel.

Sobre la solución de Fine al problema de la fundamentación

Marta Campdelacreu*
Universidad de Barcelona

Según el *co-ubicacionismo* una estatua y el trozo de arcilla del cual está hecha son dos objetos diferentes que pueden estar co-ubicados durante toda su existencia.

Uno de los principales problemas para el *co-ubicacionismo* es el *problema de la fundamentación*. Si (como parece) la estatua y el trozo de arcilla comparten su materia, sus partes, su forma, el lugar en el que se encuentran, etc., entonces ¿qué explica que la estatua sea una *estatua* y no sea un *trozo de arcilla*, y que el trozo de arcilla sea un *trozo de arcilla* y no sea una *estatua*?

En este trabajo me centraré en la solución de Kit Fine a este problema y argumentaré que no es una solución satisfactoria.

1 La solución de Fine

Fine presenta su solución en términos (aproximadamente) de una esfera de metal, s , permanentemente co-ubicada con el correspondiente trozo de metal, p . Según Fine un objeto de este tipo es un compuesto de *materia* y *forma*. La esfera de metal es un compuesto del *metal* y de la forma *forma esférica*. Si a es el metal y S la forma *forma esférica*, entonces $s=a/S$. Si a la forma *forma de trozo* le llamamos ' P ', entonces $p=a/P$.

La solución de Fine al problema de la fundamentación tiene dos partes: (1) lo que explica que s sea una esfera (de metal) es su forma S , y lo que explica que p sea un trozo (de metal) es su forma P ; (2) Fine dice:

It may also be shown that the sphere $s= a/S$ is not a *piece of [metal]* and that the piece of [metal] $p=a/P$ is not a *[metal] sphere*. For suppose that the sphere $s=a/S$ were a *piece of [metal]* [...]. It would then have to be a rigid embodiment of the form b/Q , where b is some [metal] and Q is the shape of a piece[-form]. Now there are a number of ways to show within our theory that $s=a/S$ is not identical to $q=b/Q$. (Fine 2008, p. 114)

* marta.campdelacreu@gmail.com

Después de este texto Fine muestra que su teoría determina que s no es idéntico a q . Sobre esto no voy a decir nada. Lo que voy a argumentar es que el texto muestra que Fine presupone que un objeto tiene una forma y solo una forma. Esta presuposición es esencial para su solución al problema de la fundamentación. ¿Dice Fine alguna cosa más sobre ello?

2 Formas

Fine (2017) presenta una propuesta muy elaborada sobre qué son las formas y cómo la forma de un objeto se determina. Para ello usa el caso de la forma de las fórmulas de la lógica proposicional. Según Fine (p. 514):

Given any formula, there are a number of objects each of which is a form of the formula. [...] *The form of a formula is, of course, a form of the formula, and what distinguishes it from the other forms is that it is the most fine-grained of them.*

Según Fine (2017) (algunos de los) *tipos* que un objeto ejemplifica corresponden a *formas* del objeto. Como veremos, todos los *tipos* relevantes para nuestros propósitos (*tener forma esférica* y *tener forma de trozo*) corresponden a formas de los objetos que los ejemplifican.

3 Las formas de Fine y su solución al problema de la fundamentación

Como hemos visto, la solución de Fine al problema de la fundamentación presupone que la esfera y el trozo de metal son objetos que tienen exactamente una forma. En el caso de s la forma *forma esférica* y, en el caso de p la forma *forma de trozo*. Sin embargo, como voy a argumentar, la propuesta de Fine sobre cómo la forma de un objeto se determina implica que la forma de p no es la forma *forma de trozo* sino la forma *forma esférica*. Pero entonces la respuesta de Fine al problema de la fundamentación (al menos para los casos del tipo que estamos analizando) está en claro conflicto con su teoría sobre las formas.

4 Una solución fineana

Una posible modificación de la *teoría de las formas* de Fine que nos permitiría solucionar el problema de la fundamentación de una manera cercana a su solución original es la siguiente (el tipo de propuesta no es nueva; otros autores presentan soluciones del mismo tipo): aunque la esfera, s , tiene la forma *forma esférica* y la forma *forma de trozo*, es un hecho primitivo, un hecho bruto, que la forma *forma esférica* determina la *propiedad*

clasificatoria del objeto, y que la forma *forma de trozo* no. Esto es lo que explica que *s* sea una esfera y no sea un trozo de metal (algo análogo sería el caso para el trozo de metal, *p*).

5 Sobre la solución fineana al problema de la fundamentación

En la última parte del trabajo argumentaré que esta solución fineana tiene consecuencias inaceptables. Para ello consideraré sus consecuencias para la siguiente cuestión: ¿Cómo se determina la *propiedad clasificatoria* de un objeto del tipo de *s* o *p*? Argumentaré que la respuesta a esta pregunta es análoga a la siguiente respuesta a la pregunta sobre cómo se determina la identidad en un caso de fisión: las dos personas resultantes de la fisión están relacionadas de la misma manera con la inicial pero solo una es idéntica a ella; esto es un hecho primitivo, bruto.

En el trabajo voy a argumentar que esta propuesta para los casos de fisión es insatisfactoria y que la propuesta análoga para la determinación de la *propiedad clasificatoria* de los objetos del tipo de *s* o *p* es análogamente insatisfactoria. Pero esta última se sigue de la solución fineana al problema de la fundamentación. Así, esta solución es igualmente problemática e insatisfactoria.

Referencias bibliográficas

- [1] Fine, K. (2008) “I-Kit Fine: Coincidence and Form”, *Aristotelian Society Supplementary Volume* 82: 101-118.
- [2] Fine, K. (2017) “Form*”, *The Journal of Philosophy* 114 (10): 509-535.

Superveniencia de propiedades económicas de los objetos a partir de propiedades físicas: análisis y limitaciones

Joan González
Universitat de Les Illes Balears

El análisis sobre el que se fundamenta la mayor parte de la teoría monetaria ortodoxa concibe el dinero como una cosa. Esto sigue siendo así mucho tiempo después de que el metalismo teórico, como posición ontológica sobre el dinero, haya muerto. Concebir el dinero como una cosa nos orienta definitivamente a un tratamiento ontológico del dinero basado en lo que a partir de aquí vamos a llamar un “enfoque de las propiedades”, según el cual la definición del dinero debe hacerse depender de una serie de propiedades específicas que lo capacitan especialmente para llevar a cabo sus funciones. Este clásico “enfoque de las propiedades” se halla ya en los textos clásicos sobre los cuales el análisis monetario ortodoxo ha acostumbrado a levantar los fundamentos de su comprensión sobre el dinero. Será este conjunto de propiedades específicas las que hagan a un objeto especialmente óptimo para constituirse como “concentrador” de la función “para cambiar”, deviniendo así un medio de cambio.

Los enfoques más recientes en ontología del dinero (si dejamos de lado los enfoques ficcionalistas) parecen sostener que el único análisis sólido de las “propiedades” monetarias tiene que ver con un enfoque, o bien (a) superveniente, o bien (b) emergentista de estas propiedades: se intenta una teoría por la cual las propiedades económicas de los objetos (en primer lugar, la propiedad “ser mercancía”) supervendría o emergería de algunas propiedades físicas de estos objetos. En el presente trabajo se intenta mostrar la insuficiencia de estos enfoques que intentan explicar, tanto la aparición de mercancías en primer lugar, como en segundo lugar la aparición del dinero a partir del trueque generalizado preexistente entre estas mercancías. La aplicación de modelos supervenientes o emergentistas al relato ortodoxo sobre la aparición del dinero se mueve pues en dos niveles que se superponen: la superveniencia de la propiedad de la cambiabilidad general, por un lado, y la emergencia de la concentración de la cambiabilidad sobre una sola cosa, por otro lado, como dos fenómenos supervenientes/emergentes distintos y superpuestos. Sólo el segundo fenómeno emergente lo es, en sentido estricto, de dinero. La propiedad de la cambiabilidad general es la propiedad sobre la cual se constituiría la función “para cambiar”, aplicable a todas las cosas pertenecientes a la región “cosa física”, mientras que la propiedad “ser dinero” es aquella propiedad sobre la cual se constituiría la función de

“medio de cambio”. La función de “medio de cambio” (consistente en la concentración en un solo tipo de cosa de la propiedad de la cambiabilidad general) tan solo puede instituirse sobre un tipo de cosa que presenta la capacidad de constituir la función “para cambiar” como su función propia (es decir, como la función que se explica desde el sentido histórico que le otorga su lugar en un sistema).

Para explicar la superveniencia de la cambiabilidad general podríamos olvidarnos del análisis de las propiedades del objeto “dinero” (durabilidad, transportabilidad, etc.), ya que no estamos analizando las propiedades del dinero, sino que estamos analizando una propiedad aplicable a todas las cosas pertenecientes a la región “cosa física”. La infradeterminación de la función “para cambiar” respecto a todas las cosas de la región “cosa física” nos invitaría a pensar a primera vista la superveniencia de la cambiabilidad general como una superveniencia (a) fuerte y (b) global. Mostraremos que la superveniencia de la cambiabilidad general (como propiedad económica superveniente sobre las propiedades físicas de los objetos de un mundo posible dado) no puede ser ni fuerte ni global: son perfectamente pensables mundos en los cuales a las cosas extensas, con todas sus propiedades intactas (masa, peso, volumen), no se les aplique nunca el conjunto de propiedades de las mercancías (no existan precios, no exista exposición, las cosas no sean intercambiadas); y, al mismo tiempo, son perfectamente pensables mundos en los cuales hay entidades que tienen propiedades que aplicaríamos exclusivamente a las mercancías (tener un precio, ser ex-puesta, ser intercambiable) sin que presenten ninguna de las propiedades basales propias de las cosas físicas. De estas dos posibilidades, que mostraremos, se siguen dos tesis:

- (a) La superveniencia de la cambiabilidad general como propiedad económica sobre las propiedades físicas de los objetos solo puede ser (1) débil y (2) local o regional;
- (b) El carácter débil y local de la superveniencia de la cambiabilidad general implica una inversión entre la cambiabilidad general como aplicable a mundos y el “medio de cambio” como tipo de cosa que emerge a partir de series estocásticas de intercambios en un mundo dado. Efectivamente, más que pensar en la función de “medio de cambio” como en una especialización o concreción de un fenómeno superveniente global anterior, cabría pensar si no sería el caso que la “cambiabilidad general” (la infradeterminación física de la función “para cambiar”, en el lenguaje de las funciones) supervendría como fenómeno global tan solo a partir de la emergencia de un tipo de cosa que permite el intercambio local o regionalmente; o lo que es lo mismo, si no sería el caso que la cambiabilidad general supervendría tan solo en aquellos mundos en los cuales ya ha emergido con anterioridad el dinero. De ser este el caso, cabría revisar todos los relatos sobre la génesis del dinero a partir de un trueque previo de carácter general, y con ello, buena parte de los fundamentos de la concepción ortodoxa de la teoría monetaria.

Finalizamos nuestra exposición con las consecuencias que esta inversión tendría para los fundamentos de la teoría monetaria: la posibilidad de una fundamentación filosófica a la célebre “restricción de Clower”: “el dinero compra bienes y los bienes compran dinero, pero los bienes no compran bienes”. Cabe recordar que esta restricción suponía, para Clower, “el tema central de la teoría de la economía monetaria”. Por tanto, la posibilidad de una fundamentación metafísica de esta restricción gana interés para la teoría monetaria, en el terreno de las posibilidades de su fundamentación.

Referencias

- [1] Clower, Robert: *Money and Markets*, Cambridge Univ. Press, 1986, Cambridge.
- [2] Kim, Jaegwon: *Supervenience and Mind*. Cambridge Studies in Philosophy, 1993, Cambridge.
- [3] Savellos, Elias & YALÇIN, Ümit (eds.): *Supervenience. New Essays*, Cambridge Univ. Press, 1995.

Metafísica modal contemporánea y metafísica aristotélica: ¿continuidad o ruptura?

Agustín Echavarría
Universidad de Navarra

Desde el último tercio del siglo XX asistimos a un resurgimiento de la metafísica en el ámbito de la tradición de la filosofía analítica, producido en gran medida por los desarrollos de la lógica modal cuantificada, que dio pie a diferentes propuestas de ontología para dar cuenta de la semántica de mundos posibles. Los debates en torno al realismo modal (Lewis, 1986), el proposicionalismo (Plantinga, 1974), el combinatorialismo (Armstrong, 1989), el ficcionalismo (Rosen, 1990), han ocupado un lugar preponderante en los intereses de la metafísica analítica.

Es discutido si este nuevo modo de hacer metafísica se encuentra en continuidad con la metafísica “clásica”, que va, a grandes trazos, de Aristóteles a Leibniz. Algunos autores caracterizan la renovación modal de la metafísica, sea en forma despectiva o elogiosa, como un emprendimiento de carácter pre-kantiano. Se califica de esencialismo “aristotélico” a la idea de Kripke (1980) de que nombres propios y nombres de clases naturales son designadores rígidos; Plantinga, por su parte, remite a Aristóteles y Tomás de Aquino al criticar el escepticismo modal de Quine y establecer la distinción entre modalidades *de re* y *de dicto*. Estas referencias muestran trazos de continuidad entre la metafísica modal contemporánea y la metafísica clásica de corte aristotélico.

Otros autores han subrayado rasgos de discontinuidad. Van Inwagen presenta a las modalidades (junto con la identidad, la persistencia y la mereología), como tema “propio” de la metafísica contemporánea, y señala que el interés de la filosofía clásica por las modalidades se centra exclusivamente en las modalidades lógicas *de dicto*, y no las modalidades de *de re*, estrictamente metafísicas. Se considera en ocasiones que el uso metafísico de los conceptos modales sería a tal punto un rasgo distintivo del nuevo modo de hacer metafísica, que los conceptos de necesidad y posibilidad serían los conceptos más primitivos e irreductibles, como primeros principios a partir de los cuales todos los temas de la metafísica podrían ser tratados y expresados de forma clara y precisa (Williamson, 2013).

En este contexto, ciertos intentos de desarrollar una metafísica “post-modal”, que buscan sustentar los conceptos modales en principios metafísicos anteriores y más fundamentales, como poderes causales y disposiciones (Mumford, 1998), se caracterizan habitualmente como planteamientos “neo-aristotélicos”, en oposición a la fundamentalidad que se habría concedido a los conceptos modales en las últimas décadas del s. XX.

Ahora bien, ¿es realmente la “metafísica modal” un fenómeno específico del último medio siglo? ¿Se puede, por el contrario, hablar de una “metafísica modal aristotélica” o pre-kantiana, que anticipe cierta centralidad ontológica de las nociones modales? ¿Es la metafísica de poderes una vuelta a un planteamiento genuinamente aristotélico, contraria a los desarrollos recientes de la ontología modal? El propósito de esta breve presentación es hacer un pequeña aportación para dilucidar estas preguntas. Tomando como caso testigo a Aristóteles, se intentará mostrar qué rol cumplen los conceptos modales en su planteamiento estrictamente metafísico, señalando los rasgos de continuidad y discontinuidad respecto de los planteamientos contemporáneos.

Aunque en Aristóteles los conceptos modales aparecen explícitamente tematizados en el ámbito de la lógica (*Peri Hermeneias*, c. 12), su tratamiento tiene un alcance estrictamente metafísico. Esto puede verse si atendemos al modo en que Aristóteles expone los distintos sentidos de “posible”. Por un lado, alguna cosa se dice “posible” (*dynaton*) cuando no es necesario que su contrario sea falso” (*Peri herm.* 23 7-17; *Met.* 1019b, 28-30); así, es “posible” tanto lo que es (actualmente) verdadero, como lo que puede ser verdadero, aunque actualmente no lo sea (*Met.* 1019b, 31-32). En cambio, algunas cosas se dicen “posibles” no por su oposición a la imposibilidad, sino por su relación a la “potencia” real (*dynamis*); ésta, a su vez, tiene su sentido focal en la potencia “activa”, que es el principio del cambio que se da en otro, o en lo mismo que es cambiado en tanto que otro (*Met.* 1019b, 36-1020a, 1). El primer sentido de lo posible podríamos llamarlo “lógico” o “veritativo”; el segundo es propiamente “metafísico”, al menos en dos sentidos:

- Primero, porque la potencia (y su correlativo, el acto) constituyen uno de los sentidos en que se divide el ente cuanto tal, sujeto de la ciencia primera (*Met.* 1045b, 32-33). Aristóteles señala de modo expreso que la potencia es un modo de ser real, realmente distinto del ser efectivo y actual; si no lo fuera, sería imposible dar cuenta del movimiento y la generación (*Met.* 1047a, 14-22).
- Segundo, tanto la potencia como el acto son nociones estrictamente “metafísicas” porque, a pesar de que ambas se descubren en el análisis del movimiento, sus significados van más allá de los relacionados exclusivamente con el movimiento (*Met.* 1046a, 1 y 1047a, 30); los sentidos más fundamentales de estos conceptos hacen referencia a principios constitutivos, “estáticos”, de las sustancias: la materia y la forma.

Finalmente, se puede decir que así como existe para Aristóteles una correlación entre ciertos sentidos de lo “posible” y la potencia real, también hay una correlación muy clara entre lo necesidad y actualidad (*Peri Hermeneias* 23a, 21-25; *Met.* 1015b, 9- 15). El estudio de esta correlación permite encontrar en Aristóteles los elementos que apuntan a un esbozo anticipado de metafísica modal, como el propio Estagirita sugiere de modo expreso: “[...] quizá lo necesario y lo no necesario son el principio de ser y de no ser para todo, y las demás cosas es preciso investigarlas como derivadas de éstas” (23a, 19-20).

Referencias

- [1] Armstrong, D. (1989). *A Combinatorial Theory of Possibility*. Cambridge: Cambridge University Press.
- [2] Kripke, S. (1980). *Naming and Necessity*. Oxford: Blackwell.
- [3] Lewis, D. (1986). *On the Plurality of Words*. Oxford: Wiley-Blackwell.
- [4] Mumford, S. (1998). *Dispositions*. Oxford: Clarendon Press.
- [5] Plantinga, A. (1974). *The Nature of Necessity*. Oxford: Clarendon Press.
- [6] Rosen, G. (1990). "Modal Fictionalism". *Mind* 99 (395): 327-354.
- [7] van Inwagen, P. y Sullivan, M. "Metaphysics". The Stanford Encyclopedia of Philosophy (Spring 2020 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/spr2020/entries/metaphysics/>>.
- [8] Williamson, T. (2013). *Modal Logic as Metaphysics*. Oxford: Oxford University Press.

Identidad errática y uso de la intuición

Reynner Franco
Universidad de Salamanca

La reformulación de varios presupuestos de la filosofía de la mente durante la segunda mitad del siglo pasado, especialmente a partir de la crítica al mito de lo dado, ha propiciado un punto de vista que podría dar sentido a la idea de que percibimos nuestro propio cuerpo sin que ello implique un compromiso con la creencia de que tenemos un cuerpo del que dependen en gran medida todas nuestras experiencias. La comprensión del cuerpo propio como un tipo de experiencia más, que solo puede consumarse de un modo fragmentado, se muestra como un punto de partida relevante tanto para desmitificar (destrascendentalizar o desidealizar) el cuerpo como para sugerir modos de percibirlo y su vinculación con nuestros estados conscientes.

La desidealización del yo trascendental kantiano –en la versión propuesta por Georg Simmel para reducir lo *a priori* a un ideal inductivo y sus implicaciones en la demostración de la irrelevancia de las nociones de identidad y diferencia como modos de asegurar la comprensión o articulación de lo percibido– conforma un punto de partida de esta concepción cuyo concepto articulador es el de “uso de la intuición”. Noción que parece ofrecer una respuesta básica a la pregunta sobre qué es lo comprendemos a partir de lo que percibimos del entorno, de nosotros mismos y de los otros.

Desde este enfoque *lo a priori* vinculado a la experiencia solo puede ser un resultado a posteriori donde lo que se propone como “*dado*” es solo un ideal de la experiencia, el cual no tiene por base una pretensión de validez universal, sino más bien la intensidad, continuidad, corrección, efectividad, etc. que persigue la experiencia y cuyo carácter fragmentado no puede ser trascendido, sino articulado provisionalmente desde lo que alcanzamos a comprender siguiendo la pista de un cierto “uso de la intuición”. Este sentido práctico no admite, por tanto, una prefiguración de lo sucesivo de la experiencia a partir de lo dado como premisa o desde un esquematismo, más bien lo que se considera sobrepasado desde dicho esquematismo es la percepción singular, pero esta operación no logra superar el resultado del punto de vista de la percepción: un ideal basado en la sensación y en el “uso de la intuición” que constantemente se ve afectado o alterado por cada percepción singular. El resultado más evidente es que los procesos de identificación (asimilación) o diferenciación resultan superficiales al suponer la vacua pretensión de conocer las cosas tal como son, a partir de una identidad errática.

SIMPOSIO:
PRAGMÁTICA
DE LAS RAZONES

Resumen

El eje temático del simposio une el estudio de la argumentación con el estudio de las razones, identificando las prácticas argumentativas con las prácticas en las que se piden, se dan y examinan razones. Por ello el simposio vincula la teoría de la argumentación y la epistemología, dos de las secciones (G y C) del Congreso. Hay dos maneras principales de entender la relación de la argumentación con el razonamiento. Para algunos autores la argumentación es una expresión pública de razones, mientras que para otros las razones aparecen cuando se interioriza la argumentación. El simposio parte de esta última posición, que comporta una especie de dialectificación de las razones, que permite abordar problemas de la teoría de las razones en el marco de la teoría de la argumentación y viceversa, y traducir debates de una a otra, con el resultado de un enriquecimiento mutuo.

Argumentación por analogía y paridad de razones

José Alhambra Delgado*
Universidad Autónoma de Madrid

La argumentación por analogía ha recibido bastante atención dentro y fuera de la teoría de la argumentación. Las interpretaciones coinciden en que en estas argumentaciones la semejanza entre dos elementos (foro y tema) funciona como base para asignar a uno de ellos (tema) una propiedad del otro (foro). Esta caracterización, sin ser errónea, es demasiado amplia. No evaluamos igual un argumento como (i) «el medicamento x tiene efectos beneficiosos para el organismo en ratones, por tanto es probable que también los tenga en humanos» que un argumento como (ii) «no debemos obligar a nadie a tomar el medicamento x, por las mismas razones que no obligaríamos a nadie a llevar una dieta equilibrada». Siguiendo el famoso dictum, distinguiré entre argumentos cuya razón se basa en una relación de semejanza y argumentos que hacen lo propio con una semejanza de relaciones. En (i), por ejemplo, se parte del supuesto de que ratones y humanos comparten ciertas características fisiológicas y, a partir de ahí, se argumenta que si el medicamento x tiene efectos beneficiosos en ratones, probablemente los tenga también en humanos. En (ii), en cambio, lo que se compara no son propiedades, sino relaciones; a saber, la relación entre los enunciados (a) «obligar a alguien a tomar el medicamento x supone coartar su libertad» y (b) «no se debe obligar a nadie a tomar el medicamento x», y los enunciados (c) «obligar a alguien a seguir una dieta equilibrada supone coartar su libertad» y (d) «no se debe obligar a nadie a seguir una dieta equilibrada». Mi objetivo es aclarar en qué consisten estas relaciones.

Partiré de la interpretación de los argumentos por analogía de John Woods y Brent Hudak (1989). Según estos autores cuando argumentamos por analogía comparamos relaciones de un cierto tipo; a saber, relaciones argumentativas.

«El argumento por analogía, repetimos, es un meta-argumento, un argumento en el sentido (un tanto esquematizado) de que otro argumento [...] comparte una forma idéntica con el argumento original» (Woods y Hudak, 1989: 126). En (ii), por ejemplo, se defiende que el argumento a por tanto b es aceptable/inaceptable porque el argumento c por tanto d es aceptable/inaceptable, y a es a b lo que c es a d. Woods y Hudak proponen el siguiente esquema:

1. La forma lógica de la estructura profunda del argumento A establece que las premisas de A mantienen la relación R con su conclusión.
2. El argumento B comparte con A la misma estructura profunda.

* jose.alhambra@uam.es

3. Por lo tanto, la forma lógica de la estructura profunda de B establece que sus premisas también mantienen R con su conclusión.
4. Por lo tanto, B es un análogo de A. A y B son argumentos buenos o malos, por paridad de razonamiento.

Esta interpretación es problemática. Para empezar, la forma lógica es un concepto escurridizo. Woods y Hudak aclaran que se refieren a algo más amplio que la forma veritativa-funcional tipo *modus ponens*: «una abstracción que determina la evaluación lógica para cualquier argumento que la exhiba» (Op. Cit., 1989: 128). Pero, ¿cuál es –por ejemplo– la forma lógica de un argumento como (i)? Además, tampoco queda claro en qué consisten las propiedades lógicas que se transfieren del foro al tema: ¿un buen argumento es el que nos da una buena razón para la conclusión, el que permite inferir la conclusión de las premisas, el que tiene premisas de un cierto tipo, etc.? Y finalmente, la manera de entender la relación entre la forma y las propiedades lógicas de un argumento no permite ponderar razones y, por tanto, es incompatible con la idea de fuerza argumentativa (al menos en un sentido comparativo). Este será el centro de mis críticas.

Afirmar que en un argumento por analogía se comparan argumentos es impreciso. Lo que se compara es la relación entre las premisas y la conclusión del foro y del tema, pero no el estatus de las premisas. Por ejemplo, en (ii) es indiferente en lo que respecta al argumento analógico que la premisa (c) sea verdadera o falsa, lo que interesa es que la relación entre (c) y (d) sea correcta y análoga a la relación entre (a) y (b) (esto podría explicar el carácter hipotético de algunos análogos). Algo parecido ocurre con la conclusión de los argumentos comparados. Si asumimos que el condicional «si P, entonces C» expresa (a *là* Brandom) la relación entre las premisas y la conclusión de un argumento, entonces podemos interpretar la conclusión del argumento por analogía como la afirmación del condicional asociado al tema y sus premisas como la afirmación del condicional asociado al foro, y la analogía funcionaría como la garantía toulminiana. En (ii) el argumento sería:

1. Si (c), entonces (d)
2. Por tanto, si (a), entonces (b).

Woods y Hudak interpretan estos condicionales en términos de inferencias. Para ellos «si P, entonces C» quiere decir que, si es el caso que P, entonces podemos inferir C; es decir, si P, estamos legitimados a concluir C. Esta forma de entender el condicional no tiene en cuenta la posibilidad de que haya razones para defender una conclusión contraria a C, es decir, elimina –o al menos descuida– la ponderación de razones. Propongo interpretar el condicional asociado en términos, no de inferencias, sino de razones. El condicional «Si P, entonces C» –que recordemos expresa la relación entre las premisas y la conclusión sin añadir nada al argumento– podría parafrasearse como, “P es una razón pro tanto para C”; esto es, si P, estamos legitimados a afirmar C mientras no haya razones más fuertes

para lo contrario. Así, pues, para concluir C habría que considerar – entre otras – las razones en contra, y esa es una labor que depende del contexto en el que realicemos la evaluación del argumento y no solo de sus componentes (i.e. de su forma lógica). Ciñéndonos al título del texto de Woods y Hudak, podríamos decir que un argumento por analogía es un argumento por paridad de razones.

Referencias

- [1] Woods, John, y Hudak, Brent (2020). «Por paridad de razonamiento». Traducción y presentación de José Alhambra Delgado. *Revista Iberoamericana de Argumentación*, (21), 107-115. <https://doi.org/10.15366/ria2020.21.005>.

La construcción dialógica de las razones

Hubert Marraud*
Universidad Autónoma de Madrid

El estudio de los argumentos se puede abordar desde una perspectiva inferencista o desde una perspectiva razonista. Para el inferencismo, argumentar que *A por tanto B* es aducir que *B* se sigue o puede inferirse de *A*, mientras que para el razonismo aducir que *A* es una razón para *B*. Una diferencia fundamental es que, dado *A*, el hecho de que *B* se sigue de *A*, permite directamente aseverar *B*, mientras que, en las mismas condiciones, el hecho de que *A* sea una buena razón para *B*, no autoriza sin más a aseverar *B*, puesto que puede haber razones tan buenas o mejores para lo contrario. Esta diferencia se refleja en los distintos sentidos de ‘conclusión’. Esa palabra tiene un sentido inferencista cuando se habla de la conclusión de un argumento (p.ej., *un argumento está formado por un conjunto de premisas y una conclusión*) y un sentido razonista cuando se habla de la conclusión de un diálogo o intercambio argumentativo (p.ej., *tras deliberar, llegaron a la conclusión de que...*).

¿Qué relación hay entre la conclusión de un argumento y la conclusión de un diálogo? Si se toma como básico el primer concepto, se puede concebir un intercambio argumentativo como un proceso de elección entre argumentos concurrentes, de tal manera que la conclusión del diálogo es la conclusión del argumento elegido (como en los sistemas de argumentación abstractos o las lógicas del razonamiento por defecto). Si se toma como básico el segundo, la conclusión de un argumento se define en términos del efecto de las consideraciones aducidas en la conclusión que puede extraerse en el momento en el que son aducidas. Esto nos lleva a una segunda distinción en teoría de los argumentos. Entiendo por propiedades lógicas aquellas propiedades de los argumentos que son pertinentes para responder a la pregunta «¿se debe aceptar esta proposición o propuesta por las razones aducidas para sustentarla?» y pueden describirse sin aludir ni a los efectos sobre el auditorio ni a las reglas convencionales de los intercambios argumentativos. El atomismo en teoría de los argumentos es la tesis de que las propiedades lógicas de un argumento quedan completamente determinadas por las propiedades de sus partes y las relaciones entre ellas, mientras que el holismo es la tesis de que las propiedades lógicas dependen también de elementos contextuales que no forman parte del argumento. Ir de las conclusiones de los argumentos a la conclusión del diálogo es coherente con una posición

* hubert.marraud@uam.es

inferencista y atomista, mientras que el trayecto inverso es coherente con una posición razonista y holista.

Mi propósito es mostrar cómo se construyen las razones en el curso de un intercambio argumentativo para desarrollar una teoría razonista y holista de los argumentos. Partiré de una clasificación de los diálogos según la complejidad de los argumentos que pueden aparecer en cada movimiento, propuesta por J. Anthony Blair (2012 [1998]). Mi interés por la clasificación de Blair se debe a que muestra, a través de una sucesión de diálogos cada vez más complejos, cómo se pasa de las preguntas a los argumentos y de estos a los contraargumentos. Aunque supuestamente Blair basa su clasificación en la complejidad argumentativa de los diálogos, no está claro qué se entiende por tal, a menos que asimilemos complejidad a interactividad. Por ello ofreceré una clasificación distinta de los diálogos, aunque con algunos puntos en común, basada en la complejidad de las estructuras argumentativas que aparecen en el diálogo como un todo. Esa manera de medir la complejidad de los diálogos está bien arraigada en la tradición dialéctica. Douglas Walton, por ejemplo, señala que «aunque los argumentos aparecen en un diálogo, muchas veces la mejor manera de ver un diálogo es como un argumento extenso» (Walton 1996, p.40), y James Freeman sostiene que las estructuras argumentativas deben analizarse como productos de los procesos argumentativos, y que mientras que un análisis lógico toma como unidad de análisis el argumento individual, un análisis dialéctico se centra en la argumentación completa, que posiblemente contenga múltiples argumentos que no tienen por qué formar un único argumento unificado desarrollado en el transcurso de la discusión. (Freeman 2011, p. 109).

Mi clasificación de los diálogos según la complejidad de la argumentación extendida desarrollada es razonista. Por tanto, en cada etapa del diálogo hay una conclusión provisional, que resulta de la consideración conjunta de las razones disponibles en esa etapa, y que para la tesis C puede ser C , $no-C$ o $¿C?$ A partir de esas conclusiones provisionales definiré las conclusiones de los argumentos y distinguiré seis niveles de complejidad crecientes, que corresponden al uso de encadenamientos, garantías, objeciones, recusaciones, refutaciones y ponderaciones (en la terminología de Marraud 2021).

Referencias

- [1] Blair, J. Anthony (2012 [1998]). “The Limits of the Dialogue Model of Argument”. *Argumentation* 12, pp. 325-339. Reimpreso en J.A. Blair, *Groundwork in the Theory of Argumentation*, pp. 231-244. Dordrecht: Springer, 2012.
- [2] Freeman, James B. (2011). *Argument Structure: Representation and Theory*. Dordrecht Heidelberg London & New York: Springer.

- [3] Marraud, Hubert (2021). *En buena lógica. Una introducción a la teoría de la argumentación*. Guadalajara: Universidad de Guadalajara.
- [4] Walton, Douglas N. (1996). *Argument structure, a pragmatic theory*. Toronto: University of Toronto Press.

¿Qué significa “reconocer razones”?

Paula Olmos*

Universidad Autónoma de Madrid

Los teóricos de las razones normativas manejan una definición ortodoxa de racionalidad que envolvería la “capacidad de reconocer y responder a razones” (Chang 2013, 2021; Raz 2011). Aunque semejante definición podría sugerir una pragmática de las razones, el tratamiento que se da a procesos o acciones como “reconocer” y “responder” resulta del todo ajeno a lo que parecería su lugar más obvio: las prácticas interactivas de intercambio de razones. Tan es así que la condición de razón que se reconoce (y se postula como un deber de racionalidad reconocer) a ciertos contenidos respecto de otros sería algo ya dado, con sus correspondientes notas de objetividad y facticidad (ya se trate de hechos del mundo físico o de lo que llaman hechos normativos).

Los ejemplos de reconocimiento y respuesta efectiva y exitosa son, generalmente, de carácter individual y solipsista, puramente mental en el caso de la racionalidad teórica, en la que el propio “reconocimiento” marca el paso obligado a la “creencia”, mientras que el término “responder” pretende dar cuenta de un ejercicio de la racionalidad práctico-normativa que determinará, finalmente, decisiones y acciones en el mundo físico basadas en razones.

Frente a este planteamiento y a las discusiones conceptuales a las que da lugar, partiré de la consideración de las razones como aquello que se intercambia en una interacción comunicativa para dar cumplimiento a los objetivos de la misma que, en definitiva, remiten a la cooperación entre humanos mediada por el lenguaje. En ese sentido, me adhiero a la caracterización de J. Goodwin (2016): “La idea misma de razón encierra el ser algo que aspira a ser aceptado por los demás”. Si las razones tienen sentido, surgen, se postulan como tales y en el que adquieren (o no) su condición de razón efectiva (tras su consideración y evaluación) en el intercambio, resulta que los usos “solipsistas” (consideraciones mentales, creencias y decisiones individuales) serían, en todo caso, derivados y el modelo de adquisición de la capacidad de “reconocimiento” de razones sería pragmático y parte (para muchos central) de la adquisición del lenguaje. Pero sobre todo, la capacidad de manejo de razones no se limitaría a “reconocer” y “responder” sino que, necesariamente, implica una dimensión crítica, ya que la condición de razón de un contenido para otro es una aspiración, que se dirime, precisamente, en tales prácticas. En este marco, la racionalidad no es solo “reconocer y responder a razones”, sino la capacidad interactiva de:

* paula.olmos@uam.es

dar, pedir, reconocer, examinar, contestar, criticar, debatir, rebatir, conceder, atender o responder a, etc. razones (aspiracionales).

Una vez resituada la idea del “reconocimiento de razones” en dicho marco pragmático, tendremos que asumir definitivamente su carácter abierto. Los seres humanos crecemos, habitualmente, en una comunidad de hablantes que maneja, en sus intercambios, algunos tipos relativamente comunes de razones (derivados de intercambios anteriores), discute otros, establece o trata sobre los límites de todos y, por supuesto, crea y propone nuevos tipos o usos de razones que aspiran a ser aceptados por los demás. Por ello, el “reconocimiento de razones” no se puede modelar como una mera “capacidad de acierto” sino que precisa de una habilidad para enfrentarse a novedades. Al considerar, por otro lado, como acciones o actos de los individuos en el mundo sus intervenciones en dichos intercambios, la dimensión del “responder a razones” se puede abordar (empezar a abordar) dentro de los mismos y no necesariamente como una valoración de su incidencia en acciones posteriores. En este sentido, este “responder a razones” que se da ya en el seno del intercambio comunicativo y se expresa en actos comunicativos del tipo adecuado no sería algo añadido al “reconocerlas” sino la medida en la que dicho reconocimiento, efectivamente, se da. El reconocimiento se traduce, pues, en respuesta. Y ello tanto en el caso de las razones teóricas como en el de las razones prácticas, que en esto no se distinguirían, envolviendo ambos aspectos pragmáticos de respuesta, que se concretan, finalmente, en actos (inherentemente comunicativos) de “dar”, “pedir”, “contestar”, “criticar”, “debatir”, “rebatir”, “conceder”, etc. razones.

Propongo, en primer lugar, añadir a la caracterización ortodoxa de la racionalidad en términos del “reconocer y responder a razones” precisa, las dimensiones que suponen los verbos propios del intercambio y manejo crítico de razones, y después un tratamiento integrado de todas estas dimensiones. Ya que dichas acciones son los tipos de respuesta racional que nos interesan (más allá de la puesta en práctica de decisiones racionales) y, además, son la expresión pragmática del propio reconocimiento de razones (cuya posible dimensión mental, no nos interesa aquí particularmente). Por tanto, no se trataría de dimensiones separables, como sí lo son para los teóricos de las razones normativas.

“Reconocer razones” sería, en todo caso, una habilidad práctica (un “saber cómo”) que se adquiere y se ejerce, fundamentalmente, en intercambios comunicativos y que precisa, más que de la capacidad de acierto, de la capacidad de apertura crítica. Además, como “habilidad cognitiva de propósito general”, el “reconocimiento de razones” opera en diversos contextos y, en las comunidades humanas culturalmente complejas, en una multiplicidad de prácticas que modulan y delimitan reguladamente su ejercicio. Una consecuencia importante de esta caracterización abierta y adaptativa a nuevas propuestas, a nuevos contextos y a prácticas diversas de la capacidad de “reconocer razones”, es que la distinción entre el reconocimiento que se ejerce en el seno de nuestra propia comunidad

y el que podría llegar a ejercerse (con mayores dificultades, sin duda) entre comunidades diversas no se debe entender como una barrera tajante sino, como plantea Tindale (2021), como un punto de partida para nuevos intercambios.

Referencias

- [1] Chang, Ruth (2013). “Raz on Reasons, Reason, and Rationality: On Raz’s From Normativity to Responsibility” *Jerusalem Review of Legal Studies*, December 2013, 1-21.
- [2] Chang, Ruth (2021). “What is it to be a Rational Agent?” In R. Chang y K. Sylvan, eds., *The Routledge Handbook of Practical Reason*, New York: Routledge.
- [3] Raz, Joseph (2011). *From Normativity to Responsibility*. Oxford: Clarendon Press.
- [4] Tindale, Christopher (2021). *The Anthropology of Argument. Cultural Foundations of Rhetoric and Reason*. New York: Routledge.

Pragmática de la publicidad. El “amor” por las marcas y la creación de una comunidad discursiva como estrategia argumentativa

Raúl Urbina*
Universidad de Burgos

Puede decirse que la publicidad “moderna” surgió en el ámbito de las sociedades de consumo cuando empezaron a convivir varios productos destinados a un mismo fin. En el momento en el que empezaron a existir posibles competidores, se hace necesario establecer una comunicación eficaz con los posibles compradores para que estos opten por uno de los productos. Puesto que los productos (pongamos, por ejemplo, un jabón para lavar la ropa) tienen una finalidad muy parecida, la publicidad empieza a establecer estrategias persuasivas para que el consumidor potencial se decida por el producto que un fabricante desea vender.

Paulatinamente, los mecanismos publicitarios para persuadir a los receptores fueron refinándose y los creativos empezaron a utilizar “razones” basadas en la ponderación. La ponderación del producto resulta un procedimiento muy eficaz en esa finalidad de aportar motivos de compra en los receptores, en cuanto “canta” la excelencia de un producto y, muy frecuentemente, establece un contraste valorativo con otros productos de su misma clase en los que surge un elemento positivo diferenciador (Robles Ávila, 2005).

Existe una fuerte conexión entre la creación de estereotipos (Fernández & Reyes, 2004; Fernández Gómez & Ramos-Serrano, 2010) y los recursos de ponderación. El discurso publicitario, inevitablemente, se desarrolla en un campo de juego en el que son valores decisivos la necesidad –más que en cualquier otro tipo de discurso– de la función fática (Urbina Fonturbel, 2009b, 2009a) y del acto de habla directivo que ha de permanecer siempre en el campo de lo implícito. Esto determina que sea un tipo de comunicación relativamente sencilla. No obstante, esta aparente sencillez despliega una batería argumentativa basada en los recursos léxicos y gramaticales que consiguen destacar a los productos publicitados por encima de los de su “clase”. En definitiva, la pragmática aparece en los aspectos latentes inferenciales relacionados auténticamente, mientras que la ostensión nos muestra todo un catálogo de aspectos multimodales que atraen (y distraen), frecuentemente, la atención.

* rurbina@ubu.es

En el camino hacia una persuasión publicitaria más eficaz, se percibe desde hace unos años un transvase desde las estrategias argumentativas centradas en el producto a las estrategias argumentativas centradas en la marca (Fernández Gómez & Ramos-Serrano, 2010; Nomdedeu, 2012). No se trata ya, solamente, de “vender” un bien de consumo, sino de atraer a los receptores hacia una marca. En esta tendencia publicitaria actual, se pretende que los receptores se sientan partícipes del conjunto de valores que representa una marca (Bassat, 1993)e, incluso, en el caso de algunas de las grandes estrategias publicitarias y de marketing, sean sus valedores y sus defensores. El máximo exponente de esta tendencia es el de las lovemarks (Fernández Gómez & Rodríguez Centeno, 2009; Pawle & Cooper, 2006; Roberts, 2004).

En los casos de adhesión a una marca, se percibe con más intensidad el paso hacia la argumentación publicitaria de carácter emocional (Heath, 2007; Holbrook, Morris B.O’Shaughnessy, 1984; Kemp et al., 2012). Para ello, el paso previo es de la creación de una comunidad discursiva, en la acepción que da al término Dominique Maingueneau (Charaudeau & Maingueneau, 2002; Maingueneau, 2014), vínculo entre la marca y sus valedores, que establece toda una batería de elementos argumentativos multimodales (Macagno & Botelho Wakim Souza Pinto, 2020; Marraud, 2018) que aportan una fuerza argumentativa muy eficaz en sus objetivos performativos.

Referencias

- [1] Bassat, L. (1993). *El libro rojo de la publicidad. Ideas que mueven montañas*. Folio.
- Charaudeau, P., & Maingueneau, D. (2002). Dictionnaire d’analyse du discours. En *Dictionnaire*.
- [2] Eguizábal, R. (2011). *Historia de la publicidad* (2ª). Fragua.
- [3] Fernández, E. C. G., & Reyes, I. G. (2004). Los Estereotipos de Mujer en la Publicidad Actual. En *Questiones Publicitarias* (Vol. 1, pp. 43-64).
- [4] Fernández Gómez, J. D., & Ramos-Serrano, M. (2010). Mito, publicidad y branding: un análisis mitocrítico de los sectores de la cosmética, la moda, la perfumería y el alcohol. En S. Robles Ávila & M. V. Romero Gualda (Eds.), *Publicidad y lengua española: un estudio por sectores* (pp. 54-70). Comunicación Social Ediciones y Publicaciones.
- [5] Fernández Gómez, J. D., & Rodríguez Centeno, J. C. (2009). De la Lovemark a la marca de culto: nuevos planteamientos para la gestión de las marcas. En M. I. de Salas Nestares (Ed.), *Actas del I Congreso Internacional Brand Trends: Valencia, del 18 al 20 de febrero de 2009, 2009, ISBN 978-84-692-2290-4*. CEU Universidad Cardenal Herrera. <https://idus.us.es/xmlui/bitstream/handle/11441/29578/lovemark.pdf?sequence=1>

- [6] Heath, R. (2007). Emotional Persuasion in Advertising: A Hierarchy-of-Processing Model. *University of Bath May*, 23, 2007–07. <http://opus.bath.ac.uk/287/>
- [7] Holbrook, Morris B.O'Shaughnessy, J. (1984). The Role of Emotion in Advertising. *Psychology & Marketing*, 1(2), 45-64. <https://doi.org/10.1002/mar.4220010206>
- [8] Kemp, E., Bui, M., & Chapa, S. (2012). The role of advertising in consumer emotion management. *International Journal of Advertising*, 31(3), 339-353. <https://doi.org/10.2501/IJA-31-2-339-353>
- [9] Macagno, F., & Botelho Wakim Souza Pinto, R. (2020). Reconstructing Multimodal Arguments in Advertisements: Combining Pragmatics and Argumentation Theory. *Argumentation*. <https://doi.org/10.1007/s10503-020-09525-z>
- [10] Maingueneau, D. (2014). *Discours et analyse du discours*. Armand Colin. Marraud, H. (2018). Arguments from Ostension. *Argumentation*, 32(3), 309-327. <https://doi.org/10.1007/s10503-017-9435-9>
- [11] Nomdedeu, I. (2012). El Branding Emocional: una aproximación heurística y profesional a su metodología y campos de innovación. En *Fòrum de Recerca* (Número 17, pp. 575-588). Servicio de Publicaciones. <https://dialnet.unirioja.es/servlet/articulo?codigo=4502632&info=resumen&idioma=SPA>
- [12] Pawle, J., & Cooper, P. (2006). Measuring emotion - Lovemarks, the future beyond brands. *Journal of Advertising Research*, 46(1), 38-48. <https://doi.org/10.2501/S0021849906060053>
- [13] Roberts, K. (2004). *Lovemarks. El futuro más allá de las marcas*. Empresa Activa.
- [14] Robles Ávila, S. (2005). La ponderación en el discurso publicitario. *RILCE: Revista de filología hispánica*, 21(2), 263-280. <http://dadun.unav.edu/handle/10171/6712>
- [15] Urbina Fonturbel, R. (2009a). Emociones, publicidad y retórica de las pasiones. En I. Ferreira & M. del M. Gómez Cervantes (Eds.), *Retórica e Mediatização II* (pp. 127-145). Universidade da Beira Interior. http://urbinavolant.com/ubu/wp-content/uploads/2014/06/Urbina_Emociones-y-publicidad_retorica_livroslabcom_05_09.pdf
- [16] Urbina Fonturbel, R. (2009b). Las operaciones retóricas de inventio y dispositio y la creatividad en el lenguaje publicitario. En T. Arcos Pereira, J. Fernández López, & F. Moya del Baño (Eds.), *Pectora mulcet: estudios de retórica y oratoria latinas* (p. vol. II 1291-1407). Instituto de Estudios Riojanos.

SIMPOSIO:
PRÁCTICAS
ARGUMENTATIVAS

Resumen

El simposio, que se encuadraría en la sección G del X Congreso, presenta una muestra de la diversidad de problemas que pueden plantearse en teoría de la argumentación y su carácter interdisciplinar. La teoría de la argumentación puede definirse como el estudio interdisciplinar de las prácticas argumentativas, es decir, las prácticas que consisten, total o parcialmente, en pedir, dar y recibir razones. Definir así el objeto de la teoría de la argumentación comporta una naturalización de la disciplina, que plantea cuestiones metafilosóficas, como su ubicación dentro de la filosofía y su función crítica, descriptivas, el funcionamiento de determinadas formas de argumentación y los efectos perlocutivos de la argumentación, o técnicas, cómo diagramar argumentos y argumentaciones.

Diagramación argumentativa con herramientas visuales

Enrique Alonso*
Universidad Autónoma de Madrid

Soledad Hernández**
Universidad Complutense de Madrid

Sobre la utilidad de las herramientas de diagramación en Teoría de la Argumentación

Las herramientas que se sirven de recursos visuales, ya sea en entornos de escritorio u online, para la diagramación de argumentos no se pueden considerar una novedad. En 2001 Chris Reed y Glenn Rowe presentaron la que seguramente sea la decana entre todas ellas, Araucaria, en torno a la cual se ha desarrollado una notable investigación y una no menos notable literatura científica.

Si bien este tipo de herramientas ya no resultan novedosas, lo cierto es que constituyen un recurso imprescindible en cualquier grupo de investigación maduro en el ámbito de la Teoría de la Argumentación. En este sentido, caben dos opciones: adoptar alguna de las ya existentes o desarrollar una propia. Aquí presentaremos una herramienta online, basada en Scratch, desarrollada por los autores que permite el uso de cualquier esquema de diagramación ya sea original, o de los varios existentes en la literatura.

La utilidad de este recurso, al que provisionalmente hemos denominado *Disputatio*, se reparte en dos ámbitos que resultan en cierto modo naturales: el docente y el experimental. Dado el grado de desarrollo de *Disputatio* podemos presentar avances en ambos campos con un mayor énfasis en el segundo de ellos.

Uso docente

La elección de Scratch como punto de partida de nuestro proyecto no es casual en este sentido. Este entorno de trabajo fue desarrollado por primera vez en 2003 en el Media Lab del MIT. Originalmente fue concebido como un entorno de uso docente en programación elemental para niños. Sus posibilidades se fueron explorando hasta que en 2013 se ofreció a los usuarios la opción de desarrollar sus propios componentes dentro de Scratch. Esta

* enr.alonso@gmail.com

** mahern32@ucm.es

opción nos ha permitido desarrollar, por ejemplo, un entorno de trabajo fiel al modelo de Toulmin capaz de almacenar y procesar las respuestas sugeridas, aunque no sea ese su fin inmediato. La figura 1 ofrece una idea del marco de trabajo.

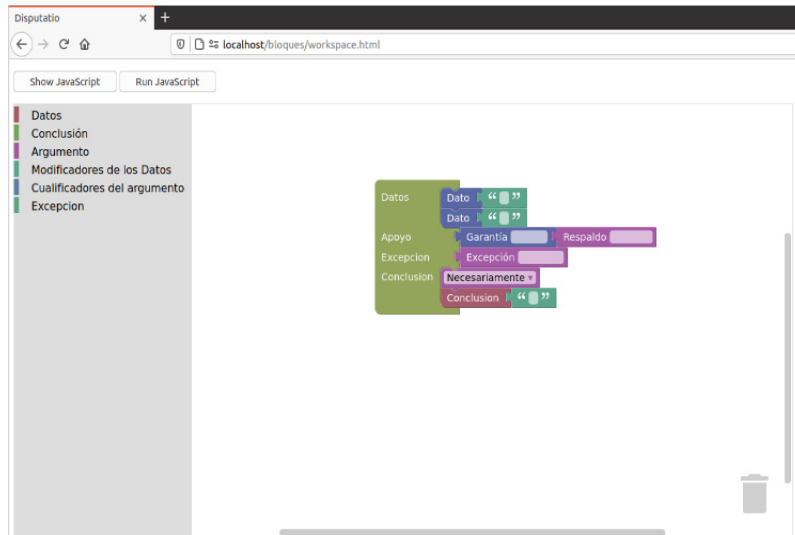


Figura 1: Ejemplo de un argumento según el modelo de Toulmin sin contenido concreto

Como es obvio, nuestra intención no es limitarnos al modelo de Toulmin, de sobra conocido, sino ampliarlo a otros modelos diagramáticos más recientes y sofisticados.

Uso en experimentación

Aunque existe un cierto trabajo hecho en el ámbito experimental en Teoría de la Argumentación, lo cierto es que es mucho menor en comparación con otras líneas dentro de la misma disciplina. Para los que mostramos una clara inclinación empírica esto constituye un problema que, solo en cierta medida, encuentra reparo en los estudios de pedagogía que han decidido adentrarse en esa vía. En definitiva, todo parece indicar que se ha dedicado mucho más tiempo y esfuerzo a determinar cómo se debería argumentar que a comprender cómo de hecho se argumenta.

Estudio de Caso 1

En conversaciones mantenidas con Fernando Leal hace algún tiempo, convinimos, gracias a su consejo, que una investigación poco atendida, pero aparentemente elemental, es aquella que analiza en qué momento los sujetos experimentales aprecian, o no, la presencia de un argumento. Con tal fin se ha diseñado un espacio de trabajo elemental con elementos ciegos -sin descripción aparente acerca de su función- tal y como se muestra en la figura 2:

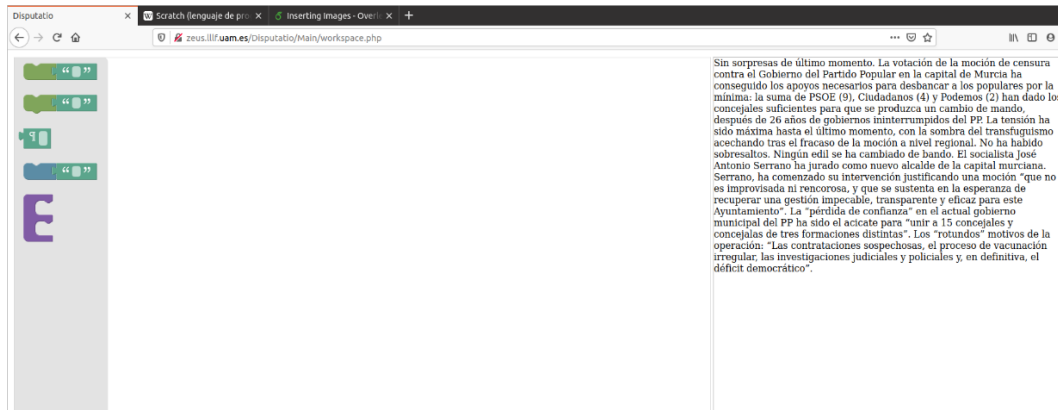


Figura 2: Ejemplo del entorno de trabajo orientado a la identificación de argumentos

Como se puede ver, en este caso se ha omitido cualquier indicación acerca de la función de las piezas disponibles. Tampoco se aclara si el sujeto se encuentra o no ante un argumento, en su lugar se le solicita que represente con los elementos disponibles el texto que aparece cargado en la parte derecha del espacio de trabajo. Algunos de los bloques disponibles disponen de espacios que sugieren la posibilidad de incorporar texto. Se informará a los sujetos que pueden hacerlo siempre y cuando se limiten a arrastrar texto desde la ventana derecha del espacio de trabajo. En otra modalidad del experimento se admite la posibilidad de incorporar texto propio.

La secuenciación del experimento se realizará por niveles educativos, empezando al menos en educación Primaria y terminando en el Ciclo superior –intentando, en este caso, que no haya habido un contacto previo con la Teoría de la Argumentación–.

Se dispone ya de un acervo de textos adaptados a los diversos niveles y que formarán el marco experimental en la fase de implementación.

La aplicación permite recoger cualquier respuesta gráfica que los sujetos sometidos a la prueba deseen ofrecer traduciéndola a un formato más tratable. El lenguaje elegido para la versión compacta de las respuestas es XML, que cuenta con una larga tradición en el ámbito de los estudios de la argumentación. Al no existir restricciones sintácticas de ningún tipo, confiamos en poder obtener un mapa detallado del modo en que los sujetos estructuran y analizan textos de diversa complejidad desde el punto de vista de la argumentación.

Estudio de Caso 2

La importancia de los marcadores argumentativos ha recibido mucha atención en la literatura científica a lo largo del tiempo. El segundo estudio de caso que estamos poniendo en marcha tiene que ver precisamente con estos elementos de los textos y contextos argumentativos. Lo que se pretende es observar en grupos experimentales debidamente seleccionados, las variaciones en la percepción de la estructura de un texto en función la

presencia o ausencia de este tipo de marcadores. Los textos serían privados de marcadores o se les incorporaría de forma explícita y se someterían, en ambas versiones, a los grupos experimentales. Una derivada de este estudio de caso sería al evaluación del peso que los marcadores pueden tener en la visualización de textos no argumentativos o dudosamente argumentativos, como tales cuando su presencia de manifiesta de forma muy expresa.

Tanto en el Caso 1 como en el Caso 2, se procederá a elaborar en detalle los protocolos experimentales que aquí solo se han esbozado de forma superficial.

Efectos perlocutivos de la argumentación

Cristina Corredor*
(UNED)

El título de esta propuesta presupone un análisis de la argumentación basado en la noción de acto de habla y la distinción locutivo-ilocutivo-perlocutivo. Desde este marco, preguntar por los efectos perlocutivos de la argumentación es preguntar por cualquier tipo de efecto que pueda considerarse causado por un acto de argumentar. Esta apreciación es demasiado indiferenciada. Es posible distinguir entre efectos perlocutivos ‘internos’, ‘inherentes’ o ‘intrínsecos’, y efectos perlocutivos ‘externos’ o ‘extrínsecos’, en correspondencia con lo que se han considerado tipos de fines de la argumentación (Mohammed, 2015). Las perlocuciones cubrirían un espectro más amplio, pues incluirían aquellos efectos que no pueden verse como fines vinculados con la actividad de argumentar. (Que no puedan verse como fines no es solo, ni fundamentalmente, una cuestión que dependa de las intenciones de quien argumenta; la argumentación es una actividad comunicativa que se realiza en un ámbito interpersonal y social, y cuáles sean los fines que quepa asociarle dependerá mucho de se practica, se reconoce y se evalúa cómo esa actividad en esos ámbitos.)

Teniendo en cuenta los que puedan reconocerse como fines de la argumentación desde la perspectiva apuntada, los efectos perlocutivos inherentes o internos serían aquellos directamente relacionados con los fines inherentes el acto de habla y, podría añadirse, son efectos en función de los cuales el acto de habla ha sido diseñado. Por ejemplo, justificar una afirmación (Johnson), convencer a un crítico razonable (pragma-dialéctica), o mostrar que una afirmación de referencia es correcta (Bermejo-Luque) pasarían a ser efectos perlocutivos internos, frente a otros efectos perlocutivos externos como lograr una posición negociadora, ver tenidos en cuenta los intereses individuales, o ser considerado vencedor en un debate político. Un ejemplo destacado de defensa de esta distinción es la pragma-dialéctica. Eemeren y Grootendorst entienden que convencer es un efecto inherente o intrínseco a la argumentación, que la define como tal. Declaran: “In our definition we have linked convincing to the performance of the illocutionary act complex argumentation” (1984, p. 4). En su definición, vinculan el acto de habla de argumentar con la condición esencial de que cuente como un intento por parte de la hablante de convencer a su oyente (ibid., p. 43). Y vinculan el efecto de convencer con la dimensión interaccional

* ccorredor@fsof.uned.es

de la argumentación, identificando la dimensión comunicativa con el efecto ilocutivo de lograr la comprensión del acto: su contenido y su fuerza.

Creo que es correcto diferenciar dos grandes tipos de efectos perlocutivos, distinguiendo el tipo de efecto que consiste en convencer, en lograr la aceptación del acto de habla, de otros tipos de efectos fortuitos, intencionados o no. Sin embargo, no comparto la intuición de que la dimensión comunicativa debe identificarse tan solo con la comprensión (uptake), identificando así dimensión interaccional con lo perlocutivo. Eemeren y Grootendorst introducen con claridad esta idea, cuando escriben: “Our hypothesis is that in the communicative sense argumentation is a form of language use corresponding to the forms of language use characterized in the speech act theory as illocutionary acts and that as regards its interactional aspects argumentation is linked with the perlocutionary act of convincing” (ibid., p. 300). A esta comprensión del acto de habla de argumentar puede oponerse una aproximación interaccionista a los actos de habla. Aquí, el efecto ilocutivo (convencional, en el sentido de Austin, 1962) se explica diciendo cómo el acto de habla ha introducido cambios en las posiciones normativas de las personas interlocutoras: en sus compromisos, responsabilidades y obligaciones, así como en sus habilitaciones, autorizaciones y derechos. Este tipo de efecto es interaccional, y no es perlocutivo. Antes bien, que el acto de habla dé lugar a esos cambios es lo que permite considerar que se ha logrado correctamente y con éxito.

Puede considerarse un ejemplo para mostrar la diferencia entre estas dos intuiciones.

“El portavoz del Gobierno madrileño, Enrique Ossorio, ha hecho un llamamiento a los habitantes de la Comunidad para que “confíen en las vacunas”, que son “seguras” y cuyos beneficios “son muchos más que los escasos riesgos” que puedan tener.” (El País, 11 de abril de 2021, edición online).

Un análisis básico de este fragmento, haciendo uso del modelo de Toulmin, permitiría identificar al menos estos elementos. La afirmación que el hablante quiere justificar es:

(C) Puede confiarse en las vacunas

Las razones que se ofrecen son dos:

(R1) Las vacunas son seguras

(R2) Los beneficios de las vacunas son muchos más que los escasos riesgos.

Cabe considerar que R1 y R2 están a su vez constituyendo un acto de argumentación subordinado, de forma que R2 está a su vez justificando (dando una razón para) R1.

Sin embargo, lo que me interesaría son los posibles efectos perlocutivos de este acto de argumentación. Quienes escuchen o lean la declaración pueden reaccionar con

aprobación, en el sentido de que acepten el acto de habla en tanto que justificativo de la afirmación; en ese caso, se habría logrado el efecto perlocutivo de convencer. La misma declaración puede provocar indiferencia, desconfianza, duda, o incluso irritación. Todos estos efectos, también perlocutivos, no serían intrínsecos o inherentes a un acto de argumentar, o sí.

Todavía teniendo en cuenta el modelo de Toulmin, lo que hace de este acto de habla un acto de argumentar es que la afirmación de referencia (C) y las razones que se ofrecen para su aceptación (R1, R2) están internamente vinculadas por una garantía: una licencia inferencial que autoriza a dar el paso desde R1 y R2 a C, y cuya reconstrucción explicitada podría ser: razones tales como R1 y R2 permiten afirmar C (cf. Toulmin, 1958, p. 134). Pero, además, para que pueda decirse que el acto argumentativo se ha realizado correctamente y ha resultado afortunado, es condición necesaria que haya tenido el efecto ilocutivo, no perlocutivo, de establecer determinadas relaciones normativas entre el hablante que argumenta y su auditorio. Esta es la hipótesis que me propongo detallar y defender en esta propuesta. Como resultado de ello, la dimensión interaccional no podrá considerarse restringida a los efectos perlocutivos internos o externos del acto de habla argumentativo.

La singular práctica de la argumentación vía *exemplum*

Luis Vega Reñón*
Profesor emérito, UNED

Aunque la argumentación por el ejemplo tiene múltiples usos y cumple relevantes servicios en la práctica discursiva cotidiana, ha representado tradicionalmente un pariente pobre en la Retórica y la Teoría de la argumentación desde Aristóteles hasta Douglas Walton. Por fortuna, el *exemplum* ha cobrado cierto relieve en nuestros días, gracias al interés de medievalistas e historiadores de la literatura y al calor del reconocimiento y la vindicación actuales de la narración argumentativa. El *exemplum* me interesa (1º) por su propia singularidad; (2º) por su capacidad para desafiar algunos tópicos arraigados como la naturaleza entrecruzada (pro-con) de la confrontación dialéctica; y (3º) por evidenciar las dificultades de reconocer e integrar teóricamente las prácticas argumentativas que discurren al margen de los usos y esquemas académicos.

1. La compleja índole del *exemplum* hace difícil una definición. Sugiero esta noción aproximativa: relato breve de un suceso referente a una experiencia humana o humanizada, con un propósito suasorio activo o preventivo y una triple función: deleitar, instruir y edificar. Un suceso es un acontecer dado cuya realidad no se plantea ni cuestiona; lo que más importa es su pertinencia para el caso: suele venir precedido de una fórmula como: «así me acontece como le aconteció a tal y cual». Estos son personajes humanos y/o humanizados. El propósito puede ser inducir a creer o hacer algo o, por el contrario, cuidarse de ello. La triple función corresponde a la atribuida tradicionalmente a la retórica: «delectare, docere, movere».
2. Pasando a una caracterización más directa del *exemplum* conviene reparar en su composición externa antes de centrarse en su constitución interna. Desde ese punto de vista, cuenta con una estructura básica compuesta por (i) presentación de los personajes en situación; (ii) anécdota o “historia” contada; (iii) moraleja y aplicación al caso. Pero es más rica e ilustrativa la estructura extendida de *exempla* elaborados como los del Libro de buen amor que comprenden: (i*) presentación de personajes; (ii*) relato del suceso, (iii*) patrón o moraleja; (iv*) aplicación al caso; (v*) réplica.
3. El *exemplum* del hortelano y la culebra (1348-1356) es arquetípico. La alcahueta Trotaconventos ofrece los servicios del Arcipreste a la monja Dña. Garoza, en cor-

* lvegar.academ@gmail.com

respondencia a la atención que ha recibido de ella: (1346 b-d). Pero la dueña, de «seso sano», contesta: «Así me acontece con tu consejo vano / como con la culebra aconteció al hortelano». Ya tenemos los personajes del pretendido enredo y del ejemplo, y la urdimbre básica: la pretensión de Trotaconventos y el rechazo de la dueña, sobre la que se urden las sucesivas tramas argumentativas, bajo la forma de ejemplos, orientadas a justificar las posiciones de una y otra. Estamos ante una argumentación práctica compleja en el marco de una discusión informal. Desde un punto de vista retórico pertenece al género deliberativo en tanto que considera una disposición o acción futura, y adopta un carácter prudencial. Este es el sucedido: Un hortelano, buen hombre, halló en enero en su huerta a la intemperie y bajo un fuerte temporal una pequeña culebra aterida de frío. Apiadado se la llevó a su casa, la cuidó y alimentó hasta que se hizo grande y lo atacó. La dueña saca como moraleja la sentencia de la historia y su aplicación: «Alégrase el malo en dar por miel veneno / <...> por piedad engaño Trotaconventos replica: «Señora», dice la vieja, «¿por qué soy vituperada? / <...>/ acontéceme como al galgo viejo que no caza nada» (1356). donde bien le advino / así derechamente a mí de ti me vino» (1354). Desde un punto de vista inferencial, el *exemplum* discurre de un caso particular a un caso particular por comparación y contraposición, y a través de sentencias referidas a una representación concreta: «Alégrase el malo...». No es una inferencia inductiva ni procede por analogía. No envuelve tanto la descripción de estados, propiedades o relaciones, como el relato de un acontecer parejo, de manera que no parece descansar en una relación de similitud ni en una similitud de relaciones, sino en una historia concreta, viva y directa. En esta misma línea no solo operan los patrones, directrices y sentencias; también abundan las expresiones cargadas emotiva y moralmente. La deliberación prudencial se ve así entretejida y animada con motivos y elementos de género epidíctico (e g. 1346d).

El *exemplum* aún reviste más interés en una perspectiva dialéctica. Aquí se puede distinguir una dimensión monológica y otra dialógica. En la primera asistimos no a pruebas sino a justificaciones de los pasos subsidiarios del propósito principal, marcados por los sucesivos *exempla*, y a su tratamiento como autojustificaciones: son apelaciones *ethóticas* las que tejen la trama; la vieja y la dueña quieren dejar sentada su buena imagen antes que establecer la conveniencia o inconveniencia de la propuesta; el *páthos* viene ligado al *éthos*. La otra dimensión, dialógica, no parece menos singular: el debate discurre por medio de *exempla* independientes contrapuestos, discursos paralelos que no llegan a cruzarse o cuestionarse entre sí. No se objeta el caso aducido, no se contraargumentan ni impugnan sus alegatos, supuestos o implicaciones. Tampoco hay al final visos de refutación: solo una réplica que protesta ante la proyección personal de la moraleja y aduce como

respuesta otro *exemplum* de autojustificación. Parece una dialéctica sin diálogo, similar a la didáctica (entre un maestro y su discípulo) o la doctrinal (entre un cristiano y un judío o un mahometano) practicadas en medios académicos. Por lo demás, no se suponen ni insinúan criterios de verdad o de verosimilitud. Lo que se trasluce es una plausibilidad que descansa en la congruencia narrativa de la historia, en la pertinencia de su aplicación al caso y en su pertenencia al acervo de la cultura popular, por remitir al folclore tradicional o a experiencias familiares.

En suma, el *exemplum* consiste en una trama retórica tejida sobre una urdimbre dialéctica, donde ambas pretenden cumplir dos funciones argumentativas principales: una ilustrativa y otra inductora o preventiva de disposiciones y acciones, que propician su capacidad de convicción y de arrastre pues, según era bien sabido, «verba docent, exempla trahunt».

Filosofía de la Lógica Ordinaria: la Lógica Informal como hija de la Filosofía del Lenguaje Ordinario

Javier Vilanova*
Universidad Complutense de Madrid

Examinaré las relaciones disciplinares entre lo que se suele llamar “Filosofía del Lenguaje Ordinario” (FLO) y lo que se conoce como “Teoría de la Argumentación”, o menos frecuentemente como “Lógica Informal” (LI), para sacar algunas directrices metodológicas y consideraciones metateóricas útiles para la LI.

Primero, repasaré las etiquetas usadas para ambas disciplinas (Filosofía del Lenguaje Informal, Pragmática de la Lengua, Nueva Retórica, Pragmadiálctica,...), examinando su origen, etimología y connotaciones, y propondré dos rótulos ya usados (FLO y LI) y dos nuevos (Filosofía del Lenguaje Común, Filosofía de la Lógica Común) como mejor medio de referirse a ella.

En segundo lugar, haré un mínimo repaso histórico de las relaciones entre ambas, centrándome en la influencia de Austin y Wittgenstein en los pioneros de la LI, tomando como ejemplo paradigmático a Toulmin y el uso en LI de conceptos y teorías de FLO (como la teoría de los actos ilocucionarios de Searle o la teoría de las máximas conversacionales de Grice).

En tercer lugar, repasaré los rasgos fundamentales de FLO y LI: fenómeno real, orientación empírica, preeminencia de los ejemplos, Importancia del contexto y otros elementos pragmáticos, introducción de la diversidad y la gradualidad, pluralidad metodológica, foco en las aplicaciones, consideración de la teoría como herramienta. Defenderé que lo que las distingue en sus respectivos campos de estudio es su aproximación pragmática (tanto en la consideración del lenguaje y la argumentación como herramientas, como en su inseparabilidad de la situación en la que operan), que surge naturalmente al pasar de la especulación a priori a la observación de ejemplos reales del fenómeno estudiado. Lo apoyaré con una analogía: la relación entre la FLO y la Filosofía Analítica clásica es como la que hay entre la LI y la Lógica Simbólica clásica.

En cuarto lugar, extraeré una aplicación metateórica para la FLO a partir de su filiación con la LI. Comenzaré señalando un equívoco en el que incurren muchos de los filósofos no familiarizados con la lógica informal y sí con la filosofía del lenguaje ordinario, y especialmente quienes siguen la estela de Austin y Wittgenstein, y que surge de identificar

* vilanova@filos.ucm.es

“lógico” y los conceptos relacionados (“demostración”, “deducción”, “válido”, “falacia”, etc.) con los de la lógica formal. Deshacer este equívoco, además de para refutar la clásica acusación a la Filosofía del Lenguaje Ordinario de que conduce al irracionalismo o incluso al escepticismo, servirá para que quienes se sienten cómodos trabajando en el marco de la FLO puedan incorporar a su discurso una dimensión argumentativa más rica y fructífera.

En quinto lugar, usaré el parentesco disciplinar para resolver la imposibilidad a la que parecen abocados los lógicos informales para ejercer, o al menos servir de ayuda, una función normativa, crítica o correctora de la propia práctica que toman como objeto de estudio. Según Woods (2000) la LI debería acometer las siguientes tareas:

- A1. Especificar las condiciones de la cogencia de un argumento y tratar de justificarlas.
- A2. Desarrollar una teoría de cada una de las propiedades impuestas por esas condiciones (por ejemplo, la relevancia, la suficiencia y la aceptabilidad).
Una consideración del más puro sentido común («me parece que las premisas nunca justifican las conclusiones. Son las personas las que justifican, y son las acciones los planes y las políticas sobre las que recae la justificación») hace que Woods cobre consciencia de que de nada vale poseer los criterios aludidos en A2, si los humanos no contamos con medios para aplicarlos, y en consecuencia añade otro requisito:
- A3. Determinar las condiciones en las que son reconocibles las propiedades y condiciones de los argumentos cogentes –es decir, en las que un sujeto competente y bien situado puede saber si la propiedad es instanciada y la condición satisfecha.

Woods solo encuentra dos medios de cumplir esta exigencia: la “toulminización” y la “dialectificación”, y ambos resultan problemáticos. La toulminización delega esta función en las disciplinas a las que pertenecen los argumentos los argumentos jurídicos en el Derecho, los geológicos en la Geología, etc.) Woods lo identifica con la estrategia de Toulmin porque en su modelo las garantías son puestas por los campos de argumentación y no por el filósofo. Si es así, concluye Woods, el filósofo pierde toda su capacidad crítica, y con ella todo interés para los especialistas del campo correspondiente. La dialectificación utiliza la propia teoría del lógico informal para plantear la pregunta de si un determinado argumento cumple sus criterios y responderla argumentando colectivamente. Aparte de la grosera circularidad metodológica que parece comportar, la dialectificación borran la frontera entre el experto en argumentación y el lego, con lo que pierde de nuevo la capacidad crítica. Woods concluye que A3 es un objetivo no cumplido por la LI, y pasa a examinar las posibilidades de una LI puramente descriptiva.

Creo que el problema surge cuando el especialista cede al usuario “la última palabra” sobre la validez de los argumentos, algo consecuente con su orientación pragmática.

Examinaré cómo resuelven la dificultad Austin y Wittgenstein (cada uno a su manera) con respecto al lenguaje, e intentaré aplicar su *modus operandi* a la LI. La clave es no concebir la LI como una “Ingeniería de la Argumentación” ni como una “Matemática de la Argumentación”, como en el viejo paradigma de la Lógica Formal, sino en una “Física de la Argumentación” que no impide que tenga aplicaciones prácticas.

Para concluir, propondré considerar la LI como una parte de la FLO, y por tanto como una investigación lingüística. Por un lado, la LI es una reflexión genérica sobre ese fenómeno tan específicamente humano de influirnos mutuamente a través del lenguaje, no con imperativos, sino con razones, que a su vez propicia un intercambio de razones que lleva a conclusiones colectivas más satisfactorias que las obtenidas individualmente. Por otro, estudia los repertorios de convenciones lingüísticas que nos permiten llevar a cabo, en distintos ámbitos y para distintos objetivos, esos procesos de exposición e intercambio de razones.

Referencias

- [1] Woods, (2000), “How philosophical is informal logic?”, *Informal Logic* Vol. 20, No.2 (2000): pp. 139-167.

SIMPOSIO:
CUERPOS SEXUADOS: PERSPECTIVAS
FILOSÓFICAS SOBRE TEORÍAS
BIOLÓGICAS DE LA DIFERENCIA
SEXUAL Y LA IDENTIDAD DE GÉNERO

Resumen del simposio

En este simposio se examinan varias áreas de investigación biológica y biomédica donde se han producido desarrollos empíricos y conceptuales importantes en torno a los cuerpos sexo-generizados, sobre los que se presentan cuatro trabajos. El primero de ellos interviene en el debate sobre el estatus metafísico del sistema Sexo/Género, proponiendo una nueva caracterización de las diferencias sexuales como clases emergentes. El segundo presenta un análisis crítico de distintas teorías e hipótesis neurobiológicas en torno a las identidades trans*, para ofrecer una reinterpretación de las mismas desde un marco relacional, procesual y múltiple. El tercero analiza aspectos filosóficos de la menstruación: su función, su normatividad y su relevancia en debates sobre salud y enfermedad. Por último, la cuarta contribución analiza críticamente distintas teorías postuladas en las últimas décadas sobre el orgasmo femenino a partir de los diversos usos y significados del concepto de homología.

Justificación del simposio

Desde los orígenes de la epistemología feminista, la reflexión crítica en torno a las ciencias biológicas ha jugado un papel central en la articulación de las teorías feministas sobre la ciencia y la tecnología. El desarrollo de la epistemología del conocimiento situado de Donna Haraway (1988) en el campo de la primatología (1989), el cuestionamiento de los roles de género en antropología (Longino 1990) o la identificación de los sesgos de género en los estudios sobre el comportamiento sexual (Hrdy 1986) son algunos de los ejemplos más sobresalientes. En la actualidad, la mirada feminista se ha consolidado en distintos campos disciplinarios, incluyendo los estudios feministas de Ciencia, Tecnología y Sociedad, la metafísica feminista (Haslanger 2018, Witt 2011) y la neurociencia feminista (Rippon 2020). Sin embargo, en el seno de la filosofía de la ciencia y, en particular, de la filosofía de la biología (Fehr 2018), el desarrollo de conceptos y teorías con perspectiva de género, incluso en ámbitos como la salud y la sexualidad, es todavía incipiente.

En esta encrucijada teórica y social, nuestra propuesta se propone explorar, desde distintas perspectivas feministas, teorizaciones biológicas y tratamientos biomédicos de la diferencia sexual, considerando cómo estos están condicionados por sesgos de género y tienen, a su vez, implicaciones sociales de radical calado. Las investigadoras reunidas en este simposio, procedentes de distintas universidades y en distintos estadios de la carrera académica, forman parte de un grupo de investigación que viene trabajando desde 2017 en los estudios feministas de las ciencias biomédicas y que en el último año ha incorporado a investigadoras que han añadido nuevas líneas de investigación. Este simposio reúne una selección de estos trabajos, en torno a cuatro proyectos en filosofía de la ciencia sobre áreas de investigación biológica y biomédica en donde se han producido importantes avances empíricos y conceptuales a lo largo de la última

década. En concreto, los trabajos incluyen teorías muy recientes sobre el desarrollo de la identidad sexo/género, la evolución del orgasmo femenino, tratamientos biomédicos para la interrupción de la menstruación, y el uso de imaginería neurobiológica en torno a las identidades trans*.

Referencias

- [1] Fehr, C., “Feminist Philosophy of Biology”, The Stanford Encyclopedia of Philosophy (2018), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/fall2018/entries/feminist-philosophy-biology/>>
- [2] Haraway, D (1988). “Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective”. *Feminist Studies*. 14 (3): 575-599.
- [3] — (1989). *Primate Visions Gender, Race, and Nature in the World of Modern Science*. New York, London: Routledge.
- [4] Haslanger, S. and Ásta, “Feminist Metaphysics”, The Stanford Encyclopedia of Philosophy (Fall 2018 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/fall2018/entries/feminism-metaphysics/>>.
- [5] Hrdy, S., (1986). “Empathy, Polyandry, and the Myth of the Coy Female,” *Feminist Approaches to Science*, Ruth Bleier (ed.) New York: Pergamon Press, 119-146.
- [6] Longino, H. E. (1990). *Science as Social Knowledge*. Princeton, NJ: Princeton University Press.
- [7] Rippon, G. (2020). *The Gendered Brain: The New Neuroscience That Shatters The Myth Of The Female Brain*. London: Vintage.
- [8] Witt, Ch. (ed.) (2011). *Feminist Metaphysics: Explorations in the Ontology of Sex, Gender & Identity*. Springer.

El sistema Sexo/Género como clase emergente: sobre la necesidad de una metafísica feminista

Laura Nuño de la Rosa*
Universidad Complutense de Madrid

Vanessa Triviño**
Universidad Rey Juan Carlos

Uno de los debates centrales en el campo de la metafísica feminista concierne al estatus ontológico del sexo y el género (Haslanger & Asta 2018, Witt 2011). Si bien tradicionalmente ambas categorías se han considerado independientes, asumiendo la naturalidad del sexo biológico y el carácter construido de la identidad de género (Beauvoir 1953), las teorías feministas de corte constructivista han cuestionado el carácter pre-dado de la diferencia sexual y su radical entrelazamiento con las categorías de género (Butler 1990). Recientemente, autoras como Anne Fausto-Sterling (2020) han abordado la naturaleza del sistema Sexo/Género desde una perspectiva interdisciplinar que trata de dar cuenta de su construcción material. En esta presentación proponemos conceptualizar el sistema Sexo/Género como clase emergente, reivindicando la necesidad de una perspectiva metafísica, y no sólo epistémica, sobre la naturaleza de la diferencia sexual y la identidad de género.

En la primera parte, ofrecemos una taxonomía de las teorías metafísicas que subyacen a las distintas conceptualizaciones de la noción de sexo y su conexión con el género. Así, prestamos atención a teorías esencialistas tradicionales que caracterizan el sexo como una clase natural cuyos miembros comparten un conjunto de propiedades esenciales. Instancian esta posición las definiciones basadas en la dotación cromosómica (para un análisis histórico-crítico de esta postura, ver Richardson 2013) o la anisogamia (Griffiths 2020). Otras propuestas más recientes sugieren conceptualizar el sexo masculino y femenino como clases históricas (Franklin 2020). Como advertíamos arriba, desde la crítica feminista a la ciencia, otras autoras se oponen a definir las diferencias sexuales humanas como categorías exclusivamente biológicas, argumentando que, al igual que ocurre con el género, el sexo no puede delimitarse al margen de criterios socio-políticos. Es más: en el ámbito filosófico más general, la proliferación de criterios filosóficos para la definición de las clases naturales y su insuficiencia para dar cuenta de las prácticas científicas de individuación ha llevado a algunos autores a rechazar la propia noción de clase natural y,

* lauranrg@gmail.com

** vanessa.trivino89@gmail.com

en particular, su aplicación al género (Dupré 1993). En este escenario, Ingo Brigandt ha argumentado recientemente que debemos abandonar la discusión metafísica en torno a las clases para adoptar una perspectiva metodológica que atienda a cómo estas se individualizan de distinta forma dependiendo de los objetivos epistémicos y/o no-epistémicos que primen en cada contexto de investigación (Brigandt 2020). En efecto, argumentaremos que el enfoque metodológico de las clases nos permite dar cuenta de cómo muchos de los desacuerdos a la hora de caracterizar la noción de sexo se derivan de los distintos propósitos (epistémicos, pero también sociales y políticos) que guían la delimitación de los sexos. Sin embargo, consideramos que esta aproximación es insuficiente. Como argumentaremos, las implicaciones metafísicas que subyacen a las distintas caracterizaciones de la clase Sexo/Género pueden ocasionar importantes conflictos pese a que los objetivos epistémicos y no epistémicos estén bien establecidos.

En la segunda parte de la presentación, ofrecemos una interpretación metafísica de la reciente caracterización del sistema Sexo/Género propuesta por Fausto-Sterling. En la reedición de su clásica obra *Cuerpos sexuados* (2020), la autora recurre explícitamente a teorías procesualistas y relacionales, caracterizando el sistema Sexo/Género como un sistema emergente que resulta de las múltiples interacciones que tienen lugar entre los individuos de una determinada cultura o sociedad. Esta propuesta difiere de las teorías disposicionales más clásicas, en tanto en cuanto considera que sólo hay un proceso por el que se establece la identidad sexual. Este proceso, que es independiente de rasgos biológicos como la dotación cromosómica o el sistema reproductivo, se bifurca en múltiples direcciones dependiendo del contexto social y de las interacciones con otros individuos. Al no tener carácter disposicional, el proceso por el que se establece la identidad sexual no depende de propiedades internas que condicionan el tipo de manifestación asociada a una disposición, sino de factores relacionales. Argumentamos que, desde una perspectiva metafísica, esta aproximación introduce una nueva conceptualización de las clases como “clases emergentes”, construidas tanto biológica como socialmente. Para concluir, analizamos críticamente una cuestión no considerada en la propuesta de Fausto-Sterling, a saber: cómo las distintas construcciones de los distintos rasgos que caracterizan al individuo, si bien no son determinantes, pueden afectar a la probabilidad de alcanzar distintos atractores sexuales.

Referencias

- [1] Beauvoir, S. (1953). *The second sex*. Translated and edited by H.M. Parshley. New York: Alfred A. Knof.
- [2] Brigandt, I. (2020). *How to Philosophically Tackle Kinds without Talking About “Natural Kinds.”* philsci-archive.pitt.edu. Retrieved from <http://philsci-archive.pitt.edu/id/eprint/17369>
- [3] Butler, J. (1990). *Gender Trouble: Feminism and the Subversion of Identity*. Routledge: New York.
- [4] Dupré, J. 1993. *The Disorder of Things: Metaphysical Foundations of the Disunity of Science*. Cambridge, MA: Harvard University Press.
- [5] Fausto-Sterling, A. (2020). *Sexing the Body: Gender Politics and the Construction of Sexuality*. New York: Basic Books.
- [6] Franklin, L. (2020). The Animal Sexes as Historical Explanatory Kinds. In *Current Controversies in Philosophy of Science*, eds. Dasgupta and Weslake.
- [7] Griffiths, P. E. (2020). Sex is real. *Aeon*.
- [8] Haslanger, S. and Ásta, “Feminist Metaphysics”, *The Stanford Encyclopedia of Philosophy* (Fall 2018 Edition), Edward N. Zalta (ed.), URL = [<https://plato.stanford.edu/archives/fall2018/entries/feminism-metaphysics/>](https://plato.stanford.edu/archives/fall2018/entries/feminism-metaphysics/).
- [9] Richardson, S. (2013) *Sex Itself: The Search for Male and Female in the Human Genome*. Chicago (Illinois): University of Chicago Press.
- [10] Witt, Ch. (ed.) (2011). *Feminist Metaphysics: Explorations in the Ontology of Sex, Gender & Identity*. Springer.

Múltiples devenires trans*: Un análisis transfeminista de las principales aportaciones de la neurobiología a la cuestión de las identidades trans*

Maite Arraiza Zabalegui*
Universidad Pública de Navarra, UPNA

Lo trans* ha ganado visibilidad en los últimos tiempos y con ello ha aumentado el interés de la neurociencia y la neurobiología por explicar y ofrecer las claves de este fenómeno. En un contexto de avances legales, insuficientes todavía a nivel global, de altos niveles de precariedad y discriminación de las personas trans*, y de álgido debate social, el análisis de las principales teorías e hipótesis neurobiológicas acerca de las identidades trans* resulta de gran relevancia, dadas sus múltiples intersecciones e implicaciones sociales.

La influyente teoría neurobiológica sobre el origen de la transexualidad explica la transexualidad como incoherencia entre el sexo genital y el sexo cerebral, que se desarrolla en la dirección opuesta, debido a perturbaciones en los niveles hormonal pre y neonatales y a factores genéticos, que fijan la identidad de género de manera irreversible (Swaab y Bao, 2013). Ello se traduce en la inversión de zonas cerebrales sexualmente dimórficas y de su número de neuronas: el BSTc (Kruijver et al., 2000; Zhou et al., 1995) y el INAH3 (García-Falgueras y Swaab, 2008), y en un número intermedio de neuronas del INAH1 (García-Falgueras et al., 2011). Diversas autoras han criticado la idea del dimorfismo sexual cerebral, también en relación a sus regiones, y señalado que la teoría neurobiológica sobre el origen de la transexualidad no demuestra el origen biológico de las identidades trans* (Eliot et al., 2021; Fausto-Sterling, 2012).

Los estudios MRI de Rametti et al. (2011a, 2011b) coinciden con la hipótesis de la teoría neurobiológica sobre el origen de la transexualidad, y afirman la masculinización de diversos fascículos de la materia blanca en los hombres transexuales y su masculinización incompleta en las mujeres transexuales. Recientemente, nuevos estudios fMRI encuentran correlatos neurobiológicos de la desconexión entre la percepción del cuerpo y la auto-identificación de las personas trans, previamente a la hormonación: menor conectividad en el cíngulo y el precúneo y en la región occipital y temporal (Feusner et al., 2017); activación de las mismas redes auto-referenciales de la corteza prefrontal

* maite.arraiza@ehu.eus

dorsomedial, pero alineadas con la identidad de género, y no con el sexo asignado en el nacimiento (Majid et al., 2020).

En esta intervención, desde una perspectiva transfeminista, problematizaremos el marco binario del que emergen los principales estudios, hipótesis y teorías neurobiológicas acerca de las personas trans*, para argumentar a favor de una interpretación basada en: (i) la hipótesis de las identidades trans* como fruto de la interacción entre los genes, las hormonas y los factores medioambientales –sociales, discursivos, tecnológicos, etc.– y que, por tanto, presentan un carácter procesual (Fausto-Sterling, 2020; Jordan-Young, 2010); (ii) una reinterpretación de los resultados de los diversos estudios neurobiológicos sobre las diferentes regiones cerebrales de las personas trans* desde la hipótesis de los múltiples mosaicos cerebrales (Alon et al., 2020; Joel et al., 2019). Finalmente, enraizaremos el concepto de disforia de género corporal en el de disforia social de género y en el imperativo social de coherencia sexo-genérica (Galofre y Missé, 2015; Walsh y Einstein, 2020), desde el que realizaremos una lectura de la hipótesis de la conectividad de las redes cerebrales implicadas en la percepción del yo y del propio cuerpo, ofreciendo así una interpretación relacional las identidades y corporalidades trans*.

Referencias

- [1] Alon, N., Meilijson, I., & Joel, D. (2020). Testing the masculinization hypothesis in a sample of 23,935 human brains. *bioRxiv*. doi: 10.1101/2020.11.09.373258
- [2] Eliot, L., Ahmed, A., Khan, H., & Patel, J. (2021). Dump the “dimorphism”: Comprehensive synthesis of human brain studies reveals few male-female differences beyond size. *Neuroscience & Biobehavioral Reviews*, 125, 667-697. doi: 10.1016/j.neubiorev.2021.02.026
- [3] Fausto-Sterling, A. (2012). *Sex/gender: Biology in a social world*. New York: Routledge.
- [4] — (2020). *Sexing the body. Gender politics and the construction of sexuality*. New York: Basic Books.
- [5] Feusner, J. D., Lidström, A., Moody, T. D., Dhejne, C, Bookheimer, S. Y, & Savic I. (2017). Intrinsic network connectivity and own body perception in gender dysphoria. *Brain Imaging and Behavior*, 11(4), 964-976. doi: 10.1007/s11682-016-9578-6
- [6] Galofre, P. y Missé, M. (Eds.) (2015). *Políticas Trans. Una antología de textos desde los estudios trans norteamericanos*. Barcelona: Egales.
- [7] García-Falgueras, A., & Swaab, D. F. (2008). A sex difference in the hypothalamic uncinate nucleus: relationship to gender identity. *Brain*, 131(12), 3132-3146. doi:10.1093/brain/awn276

- [8] García-Falgueras, A., Ligtenberg, L., Kruijver, F. P. M., & Swaab, D. F. (2011). Galanin neurons in the intermediate nucleus (InM) of the human hypothalamus in relation to sex, age, and gender identity. *The Journal of Comparative Neurology*, *519*(15), 3061–3084. doi:10.1002/cne.22666
- [9] Joel, D., Garcia-Falgueras, A., & Swaab, D. (2019). The Complex Relationships between Sex and the Brain. *The Neuroscientist*, *107385841986729*. doi:10.1177/1073858419867298
- [10] Jordan-Young, R. M. (2010). *Brain storm. The flaws in the science of sex differences*. Cambridge, MA: Harvard University Press.
- [11] Kruijver, F. P. M., Zhou J. N., Pool, C. W., Hofman, M. A., Gooren, L. J. G., & Swaab, D. F. (2000). Male-to-Female trans*sexuals have female neuron numbers in a limbic nucleus. *The Journal of Clinical Endocrinology & Metabolism*, *85*(5), 2034-2041. doi:10.1210/jcem.85.5.6564
- [12] Majid, D. S. A., Burke, S. M., Manzouri, A., Moody, T. D., Dhejne, C., Feusner, J. D., & Savic, I. (2020). Neural Systems for Own-body Processing Align with Gender Identity Rather Than Birth-assigned Sex. *Cerebral Cortex*, *30*(5), 2897-2909. doi:10.1093/cercor/bhz282
- [13] Rametti, G., Carrillo, B., Gómez-Gil, E., Junque, C., Zubiarre-Elorza, L., Segovia, S., Gómez, A., & Guillamon, A. (2011a). The microstructure of white matter in male to female transsexuals before cross-sex hormonal treatment. A DTI study. *Journal of Psychiatric Research*, *45*(7), 949-954. doi:10.1016/j.jpsychires.2010.11.007
- [14] — (2011b). White matter microstructure in female to male transsexuals before cross-sex hormonal treatment. A diffusion tensor imaging study. *Journal of Psychiatric Research*, *45*(2), 199–204. doi:10.1016/j.jpsychires.2010.05.006
- [15] Swaab, D. & Bao, A.-M. (2013). Sexual Differentiation of the Human Brain in Relation to Gender-Identity, Sexual Orientation, and Neuropsychiatric Disorders. In D. W. Pfaff (Ed.), *Neuroscience in the 21st Century. From Basic to Clinical* (pp. 2973–2998). Nueva York: Springer. doi:10.1007/978-1-4614-1997-6_115
- [16] Walsh, R. & Einstein, G. (2020). Transgender embodiment: a feminist, situated neuroscience perspective. *Journal of the International Network for Sexual Ethics & Politics, (Special Issue 2020)*, 56-70. doi: 10.3224/insep.si2020.04
- [17] Zhou, J. N., Hofman, M. A., Gooren, L. J. & Swaab, D. F. (1995). A sex difference in the human brain and its relation to trans*sexuality. *Nature*, *378*(6552), 68-70. doi:10.1038/378068a0

Por qué quiero sangrar (y por qué tú también deberías): Función, normatividad y salud menstrual

Ainhoa Rodríguez*
Universidad del País Vasco, UPV/EHU

Arantza Etxeberria**
Universidad del País Vasco, UPV/EHU

Numerosas pensadoras feministas contemplan la “mujer” desde una normatividad transitoria y menstrual que va desde la menarquia a la menopausia. Subrayando cada vez más su relevancia, la menstruación gana peso en debates filosóficos sobre salud y enfermedad en mujeres. Este trabajo revisa cuestiones unidas al análisis filosófico de la menstruación como su función, su normatividad y su relación con la salud.

La falta de consenso a la hora de ofrecer una definición de función biológica, sin embargo, ha dificultado este debate. Existen, principalmente, posiciones de carácter etiológico-evolutivo, representadas por Wright o Millikan, y disposicional-causal, presentes en Boorse o Cummins. La función de la menstruación ha sido analizada desde la etiología, sobre todo tras describir la menstruación como un mecanismo de defensa contra los patógenos de la microbiota seminal (Profet, 1993). Esta idea de función buscaría la razón por la que la selección natural habría hecho que dicho rasgo perdurase. Sin éxito en esa búsqueda, contribuciones posteriores identificaron la menstruación como un subproducto, bien del cese de la actividad endometrial con efecto de ahorro de energía (Strassmann, 1996), o de la deciduización endometrial espontánea (Finn, 1998). La menstruación fue definida como un efecto secundario, sin función selectiva y, por consiguiente, dispensable. Inadvertidamente, esta idea provocó un cambio metafísico, influyendo en la percepción de la relevancia del ciclo menstrual y comprometiendo su investigación y reconocimiento social.

Aun así, una carencia de función selectiva no debería considerarse sinónimo de una carencia de función total. Definiciones alternativas como la de Cummins subrayan la relevancia funcional de un rasgo según la contribución causal, sin ser necesariamente producto selectivo. Siguiendo esta idea, Howes (2010) destaca el papel de la menstruación contra las infecciones de transmisión sexual y, como resultado, en la respuesta inmune de las mujeres. Además, aventura que la menstruación podría cumplir funciones adicionales

* ainhoarodriguezmuguruza@gmail.com

** arantza.etxeberrria@ehu.eus

y cruciales para el organismo. Así, cabe argumentar que, sin tener una función selectiva, la menstruación cumple una función normativa indispensable a la hora de garantizar la salud de la mujer. Este argumento se apoya en un análisis de las consecuencias de la ausencia de menstruación, reflejadas en patologías como la amenorrea hipotalámica funcional (AHF) en las que la incidencia de osteoporosis, enfermedad de la arteria coronaria o la hipercortisolemia es alta.

Pese a las consecuencias de la supresión de la función ovárica y su diagnóstico en más de 17.4 millones de mujeres en el mundo (Shufelt et al, 2017), la definición de la menstruación como un subproducto ha caricaturizado la gravedad del diagnóstico de AHF, llevando a profesionales de la salud a tildarlo de adaptativo y considerarlo como un algo ventajoso en la prevención del cáncer de mama (Strassmann, 1999). Tendencias médicas recientes se han apoyado en esta idea y han avalado la supresión deliberada de la menstruación a través de anticonceptivos hormonales, cuyos efectos secundarios son pasados por alto en aras de prácticas negligentes con la salud menstrual de las mujeres. Estas tendencias, además, han sido acompañadas por intentos de medicalizar la menstruación, exagerando una posible relación entre una mayor regularidad menstrual y un aumento del riesgo de cáncer de mama (Lippman, 2004; Loshny, 2004) cuya única solución pasaría, de nuevo, por el consumo a largo plazo de anticonceptivos hormonales.

Este trabajo considera, por consiguiente, indispensable el análisis de las definiciones de función biológica ofrecidas por la filosofía, con el fin de ofrecer un concepto de salud que reconozca la relevancia de la menstruación, no sólo en lo pertinente a la reproducción, sino unida a su función complementaria en otros sistemas de nuestro organismo. Este análisis pasa por entender la salud menstrual como una dimensión dinámica y relacional que varía según la etapa vital de las mujeres y su situación. Profundizar en la salud menstrual nos obliga a tener en cuenta ciclos hormonales basados en respuestas endocrinológicas que son desencadenadas por tanto interacciones con otros organismos (parejas sexuales, microbiota, etc.), como con varios factores medioambientales. Así, en consonancia con los esfuerzos filosóficos que tratan de desarrollar una noción naturalizada de salud que pueda referirse a ella no como un estado estático sino como trayectoria en desarrollo, la menstruación se caracterizaría como un proceso en el que la temporalidad en la que la relación del organismo con el entorno se da lugar es clave.

Referencias

- [1] Critchley, H. O. D, et al. (2020), “Menstruation: Science and Society”, *American Journal of Obstetrics*, 223 (5) 624-664 .
- [2] Clough, S. (2002), “What is menstruation for? On the projectability of functional predicates in menstruation research” *Studies in History and Philosophy of Biological and Biomedical Sciences*, 33: 719-732.

- [3] Finn, C.A. (1996), "Why do women menstruate? Historical and evolutionary review", *European Journal of Obstetrics and Gynecology and Reproductive Biology*, 70 (1), 3-8.
- [4] Hitchcock, C.L (2008), "Elements of the Menstrual Suppression Debate", *Health Care for Women International*, 29, 702-719.
- [5] Howes M. (2010), "Menstrual Function, Menstrual Suppression, and the Immunology of the Human Female Reproductive Tract", *Perspectives in Biology of Medicine*, 53 (1), 16-30.
- [6] Jasienska, G. (2013), *The Fragile Wisdom: An Evolutionary View on Women's Biology and Health*, Harvard University Press: London.
- [7] Lippman A. (2004), "Women's Cycles Up for Sale", *Canadian Women's Health Networks*, 6/7 (4/1).
- [9] Loshny H. (2004), "From Birth Control to Menstrual Control. The Launch of the Extended Oral Contraceptive Seasonale" *Canadian Woman Studies/Les Cahiers de la Femme*, 24;1.
- [10] Profet, M. (1993), "Menstruation as a Defense Against Pathogens Transported by Sperm" in *The Quarterly Review of Biology*, 68 (3), 335-386.
- [11] Shufelt, C. L., Torbati, T., Dutra, E. (2007), "Hypothalamic Amenorrhea and the Long-Term Health Consequences", *Seminars of Reproductive Medicine*, 35 (3), 256-262.
- [12] Strassmann, B. I. (1996). The evolution of endometrial cycles and menstruation. *The Quarterly review of biology*, 71(2), 181-220.
- [13] Strassmann, B. I. (1999), "Menstrual Cycling and Breast Cancer: An Evolutionary Perspective". *Journal of Women's Health*, 8 (2), 193-202.

El orgasmo femenino y el concepto de homología

Silvia Basanta*
Universidad de Viena

Laura Nuño de la Rosa**
Universidad Complutense de Madrid, UCM

El orgasmo femenino constituye un caso de estudio fascinante para investigar los sesgos de género en las ciencias biológicas y en particular en la biología evolutiva (Lloyd 2006), así como las consecuencias de las teorías en torno a la reproducción femenina para los debates actuales sobre ciencia y género. En nuestro trabajo llevamos a cabo un análisis de las distintas teorías postuladas en las últimas décadas sobre el orgasmo femenino a partir de los diversos usos y significados del concepto de homología. Para ello, prestaremos atención tanto a los usos más ambiguos del término homología, presentes, por ejemplo, en los discursos de Sigmund Freud (1905) o Donald Symons (1979) sobre orgasmo femenino, como a los usos más técnicos en el ámbito de la biología, desde el concepto de homología del desarrollo al de homología funcional o fisiológica. El uso de este concepto constituye un hilo de continuidad en la representación que las ciencias biológicas han hecho del orgasmo femenino a lo largo de los últimos siglos. Argumentamos que algunos de los primeros usos de este concepto para explicar el orgasmo femenino han reaparecido en los últimos años, aún con modificaciones y en sus formas modernizadas, en el seno de la biología evolutiva del desarrollo. En particular, analizaremos las discusiones sobre el origen evolutivo del orgasmo femenino Stephen Jay Gould (1873, 1993) y, más recientemente, Mihaela Pavlivec y Gunter Wagner (Pavličev & Wagner 2016, Pavličev et al. 2019, Wagner 2016) han traído a debate. Razonamos que el olvido y la recuperación (aunque no reivindicada) de las dos visiones clásicas sobre el orgasmo femenino, a saber, la que recurría al modelo de las homologías entre la anatomía masculina y femenina, y la que ligaba el orgasmo femenino al proceso de ovulación, tiene que ver tanto con factores históricos y políticos como con el declive del pensamiento de las homologías y el auge del pensamiento adaptacionista y el desarrollo posterior del concepto de homología y las posibilidades que ha abierto. No por causalidad, Thomas Laqueur (Gallagher & Laqueur 1987; Laqueur 1987, 1992) elegiría el caso del orgasmo femenino para ilustrar las exigencias sociológicas y culturales que dejaron atrás el concepto de homología para pensar la relación entre el cuerpo masculino y femenino entre los siglos dieciocho y diecinueve. Si bien consideramos que los determinantes históricos que condujeron al abandono de

* sbasanta045@gmail.com

** lauranrg@gmail.com

las dos explicaciones clásicas del orgasmo femenino con la llegada de la modernidad han sido suficientemente explorados, en este trabajo nos ocupamos de pensar por qué estas explicaciones desaparecieron durante casi dos siglos y qué condiciones de posibilidad han permitido que hayan vuelto a ser foco de controversia en el campo biológico contemporáneo. Por último, exploraremos las consecuencias de las teorías actuales sobre el orgasmo para la conceptualización de la sexualidad femenina. Como veremos, a diferencia de las explicaciones clásicas y las hipótesis adaptacionistas, los nuevos usos del concepto de homología traen consigo la innovación de desligar la representación científica del placer sexual de las mujeres de sus funciones reproductivas.

Referencias

- [1] Freud, S. (1953). "Three Essays on the Theory of Sexuality" (1905). In: *Standard Edition of the complete works of Sigmund Freud, v. VII*, translated by James Strachey et al. London: The Hogarth Press.
- [2] Gallagher, C., & Laqueur, T. W. (Eds.). (1987). *The making of the modern body: Sexuality and society in the nineteenth century*. Berkeley, CA: Univ of California Press.
- [3] Gould, S.J. (1987). Freudian slip. *Natural History*, 96(2), 14.
- [4] Gould, S.J. (1993). Male Nipples and Clitoral Ripples. *Columbia: A Journal of Literature and Art*, 80-96.
- [5] Laqueur, T.W. (1986). Orgasm, generation, and the politics of reproductive biology. *Representations*, 14, 1-41.
- [6] Laqueur, T. W. (1992). *Making sex: Body and gender from the Greeks to Freud*. Cambridge: Harvard University Press.
- [7] Lloyd, E. (2006). *The Case of the Female Orgasm*. Cambridge, Mass.: Harvard University Press.
- [8] Pavličev, M., & Wagner, G. (2016). The evolutionary origin of female orgasm. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, 326(6), 326-337.
- [9] Pavlicev, M., Zupan, A. M., Barry, A., Walters, S., Milano, K. M., Kliman, H. J., & Wagner, G. P. (2019). An experimental test of the ovulatory homolog model of female orgasm. *Proceedings of the National Academy of Sciences*, 116(41), 20267-20273.
- [10] Symons, D. (1979). *The evolution of human sexuality*. New York: Oxford University.
- [11] Wagner, G. P., & Pavličev, M. (2016). What the evolution of female orgasm teaches us. *J. Exp. Zool. B Mol. Dev. Evol*, 326, 325.

SIMPOSIO:
ONTOLOGÍAS HÍBRIDAS:
OBJETOS CIENTÍFICOS Y CULTURAS
VISUALES ENTRE LA INDUSTRIA,
LA CLÍNICA Y EL LABORATORIO

Interés del simposio

Son muchos los materiales en la historia de la ciencia contemporánea que han circulado entre la industria, la clínica y el laboratorio. Esta propuesta tiene como objeto estudiar ontologías híbridas que comparten conceptos, representaciones, artefactos e instalaciones que conforman una cultura material de ciencia y de la tecnología en la que también participa el género.

Inserta en el marco conceptual de las historias de los artefactos científicos, esta propuesta de simposio se articula en torno a la circulación de conocimientos y objetos entre diferentes espacios: entre los escenarios de la invención y aquellos donde se utilizaron; entre el laboratorio y la clínica, la planta piloto, la fábrica, y el campo. Al hacerlo, la historia de las ciencias incluye a aquellos procesos que, mediante la adopción y el uso de experimentos, conocimientos y dispositivos, produjeron experiencias y comunidades expertas. Estas ontologías híbridas tienen en cuenta objetos y conocimientos científicos que pertenecen a distintos espacios no siempre reconocidos como profesionales. Historizar y reflexionar sobre esa circulación entre unos espacios y otros, significa también conectar colectivos de personas –mujeres, hombres, instituciones médicas, científicas e industriales– con conocimiento y género que, representado por imágenes y nuevos dispositivos –aparatos, procedimientos y métodos– dota a la historia de la ciencia y de la tecnología de una nueva narrativa.

En este simposio queremos ofrecer un conjunto de colaboraciones que tienen en cuenta dos o más de estos agentes, para brindar historias de la circulación del conocimiento entre geografías políticas, espacios profesionales y dominios de autoridad en la era contemporánea.

Significados y materialidades de un objeto híbrido: el uranio español en la década de los sesenta

Ana Romero de Pablos*
Instituto de Filosofía, CSIC

En 1966, 137 toneladas de uranio español llegaron al puerto de Nueva Orleans desde el puerto de Cádiz (España). Este viaje que puso de acuerdo las capacidades técnicas de agentes políticos, industriales y comerciales, sirve para explorar el mineral de uranio como un objeto híbrido -físico, técnico y diplomático-. Este material conectó los intereses de las autoridades franquistas, con los de los científicos y empresarios de la industria eléctrica. Sirvió al régimen de Franco para formar a técnicos e investigadores en la compra e importación de tecnologías, métodos y prácticas experimentales, y también para la introducción de formas de popularizar las políticas e ideologías atómicas norte americanas. Al tiempo que el uranio alimentaba los reactores nucleares, también creaba nuevos espacios disciplinares, modificaba paisajes y perfilaba nuevas cartografías industriales y administrativas. Un viaje en el que el uranio alteró su materialidad y también sus significados, del que participaron distintas comunidades expertas que conectaron las minas con instalaciones industriales y convirtieron este mineral en un objeto diplomático.

* ana.romero@cchs.csic.es

Imágenes digitales del cerebro en la década de 1970 y la construcción de una tecnología híbrida

Miguel Ángel Rego Robles*
Universidad de Salamanca

En 1972, en un hospital de Londres un tumor cerebral en una paciente llevó a explorar la tomografía computarizada (TC), una tecnología basada en la combinación de radiología y computación. La cabeza del paciente se introdujo en un escáner fabricado por EMI, la empresa británica líder en la producción de discos musicales. El dispositivo emitió rayos X a la cabeza del paciente durante nueve horas. En el lado opuesto estaban los receptores de la información numérica relacionada con la cantidad de rayos X absorbidos por el cerebro. Esta información fue procesada mediante un algoritmo algebraico en un laboratorio de Londres. Cada valor se recogió con una intensidad tonal diferente en blanco y negro. La matriz digital que se obtuvo se materializó en papel prensa, un tubo de rayos catódicos y una fotografía Polaroid.

El conocimiento que llevó a la construcción del dispositivo circuló entre la clínica, la industria y el laboratorio. En la construcción de la TC participaron una diversidad de agentes, incluida la computadora, la experiencia en rayos X, ingenieros electrónicos, neurorradiólogos del Hospital Atkinson Marley y el escáner de rayos X de la empresa EMI. En esta presentación se mostrarán los procesos de producción de imágenes digitales de TC que la convertirían, a partir de la década de 1970, en una tecnología habitual para las neurociencias.

* miguelangelrego@usal.es

Cocinas, bancos de semillas y laboratorios: los chiles como objetos científicos

Daniela Sclavo*
University of Cambridge

El chile se ha cosechado en México desde hace seis mil años, lo que ha llevado a una diversificación significativa de la especie. A pesar de ser el condimento más importante del mundo no ha sido estudiado por la historia de la agricultura ni de la tecnología. Tampoco ha sido analizado como un objeto científico que forma parte de una red de circulación de conocimiento. Desde las décadas de 1970 y 1980 el chile y sus variedades ha sido estudiado por genetistas, etnobiólogos y científicos conservacionistas, lo que ha llevado a que el chile y sus semillas hayan viajado desde el campo hasta los laboratorios y bancos internacionales de semillas.

Este trabajo explora cómo el chile y sus semillas han viajado en diferentes entornos profesionales y culturales. Desde las mujeres cocineras en comunidades locales a los bancos comunitarios que engloban el trabajo de la Organización para la Agricultura y la Alimentación (FAO); desde los científicos y pobladores locales, al trabajo de genetistas y etnobiólogos, hasta llegar a los grandes bancos internacionales de semillas. Esta presentación sugiere que los cultivos son parte de la historia de la circulación de los objetos científicos al mostrarlos como materiales que viajan y permiten el intercambio, adopción e integración de conocimiento en diferentes comunidades expertas transnacionales.

* ds936@cam.ac.uk

Redes colaborativas y objetos híbridos: mujeres genetistas en México de 1960 a 1970

Ana Barahona*

Universidad Nacional Autónoma de México (UNAM)

En esta presentación se abordarán las trayectorias de dos figuras destacadas de la citogenética en México, la médica argentina Susana Kofman y la médica mexicana Leonor Buentello. Ambas compartieron agenda médica y de investigación además de tiempo y lugar. Trabajaron juntas en la primera Unidad de Genética Humana del Instituto Mexicano del Seguro Social, participaron en el diagnóstico temprano de enfermedades genéticas y revelaron la correlación entre las observaciones clínicas y el cariotipo. La presentación se centrará por un lado en los cromosomas como objetos científicos híbridos que circularon entre la clínica y el laboratorio, entre la cama y el banco; y por otro, en el lugar de trabajo (la Unidad) concebido como un espacio, tanto literal como figurativo, de mujeres y hombres involucrados en la creación, difusión, apropiación y consumo del conocimiento científico. Se abordará así los contextos locales, las culturas materiales y las prácticas específicas que permitieron a estas mujeres citogenetistas mexicanas ser parte de la producción y transmisión de conocimiento en los años sesenta y setenta, gracias a su pertenencia a redes científicas colaborativas nacionales e internacionales.

* ana.barahona@ciencias.unam.mx

El feto humano como ontología híbrida: cromosomas, ecografías y fotografías en la década de 1960

María Jesús Santesmases*
Instituto de Filosofía, CSIC

Desde 1965, las fotografías del desarrollo de fetos humanos en el útero materno comenzaron a circular entre públicos expertos y legos, acompañadas de las de cromosomas fetales y de las proporcionadas por la ecografía, conjunto de imágenes fetales que creó nuevas culturas biomédicas. Esta presentación sitúa la cultura del feto público, por un lado, en las políticas de los cuerpos de la dictadura franquista y, por el otro, en el conjunto de dispositivos y significados biomédicos que manufacturaron esa cultura. Se propone ese imaginario fetal como una ontología biomédica híbrida que ha quedado inserta en las culturas contemporáneas del embarazo y del nacimiento, en las que los cuerpos de las mujeres embarazadas y sus derechos están por completo ausentes.

* mariaj.santesmases@cchs.csic.es

SIMPOSIO:
LENGUAJE Y HUMOR

Justificación del interés

A causa de un prejuicio de clara raíz “intelectualista”, el humor ha sido históricamente desatendido como objeto de estudios “serios” en la ciencia y la filosofía occidentales. Afortunadamente, desde finales del siglo XX, tanto el humor como otros fenómenos emparentados (principalmente la comicidad y la risa) han suscitado un repentino y cada vez más intenso interés, tanto en biología evolutiva (donde la risa tiende a tomarse como un criterio de demarcación para identificar el surgimiento de “lo humano” con temas tan interesantes como el papel del control de la respiración en la risa o el papel de ésta en los procesos de apareamiento y el éxito reproductivo –Provine 2016), como en neurociencia (donde encontramos un interesante nudo de problemas, desde el mínimo control voluntario de la risa hasta el llamado “punctuation effect”, que coloca la risa dentro de la conversación –Provine 2016–), así como en Sociología (donde la risa ha sido valorada tradicionalmente como un instrumento de liberación de tensiones –Berger 1997- y, más recientemente, como un medio de comunicación –Provine 2001–) y, de manera acusada, en lingüística (donde a partir de la teoría de Raskin 1981, que incluía la comprensión del humor verbal como un criterio de competencia lingüística, en el sentido de Chomsky, cada vez son más abundantes las teorías semánticas o pragmáticas del humor –Attardo 2001-). En esta mesa, haremos un repaso al tema del humor desde un punto de vista filosófico, examinando sus relaciones con el lenguaje, y, más concretamente, con el lenguaje ordinario. Para ello revisaremos el uso que hacen del humor en la investigación filosófica del lenguaje dos figuras fundadoras como Wittgenstein y Austin, exploraremos las diferencias culturales entre sentidos del humor asociadas a la diversidad lingüística, y replantaremos el papel de la risa como un logro cognitivo asociado a la comprensión profunda de la gramática, y de lo cómico como una desviación de la regularidad y a la vez denuncia de la mecanización en la práctica lingüística.

Referencias

- [1] Attardo, Salvatore (2001). *Humorous Texts: A semantic and pragmatic analysis*, Mouton De Gruyter.
- [2] Berger, Peter (1997). *Redeeming Laughter: The Comic Dimension of Human Experience*, De Gruyter.
- [3] Provine, Robert R. (2001). *Laughter: a scientific investigation*, Penguin.
- [4] Provine, Robert R. (2016). “Laughter as a scientific problem: An adventure in sidewalk neuroscience”, *Journal of Comparative Neurology* 524(8):1532-9.
- [5] Raskin, Victor (1985). *Semantic Mechanisms of Humor*, D. Reidel Publishing Company.

Resumen del simposio

Bajo la interrelación humor y lenguaje, se discutirán los usos y el funcionamiento del humor verbal para analizar su naturaleza lingüística desde:

- El valor del humor como instrumento epistémico, su vinculación con la emoción y la propuesta de una teoría sobre la risa desde los aspectos profundos de nuestras reglas lingüísticas en “El valor epistemológico del humor, la risa como logro cognitivo”.
- Los problemas de la relatividad del humor, su diversificación y los límites de la traducibilidad, así como una posible teoría de los infortunios para el humor en “Relativamente gracioso”.
- El humor como acto de habla, profundizando en el par dicotómico serio/no serio y en la noción de “humor performativo” en “La contestación de las dicotomías: La sonrisa de Austin”.
- Los procesos abductivos/interpretativos de la risa para poder entender la recepción en los juegos del lenguaje cómico y humorístico en “Siempre nos quedará la risa”.

El valor epistemológico del humor verbal: la risa como logro cognitivo

Javier Vilanova*
Universidad Complutense de Madrid

En esta comunicación pretendo reivindicar el valor del humor como instrumento epistémico basándome en la conexión que de manera constitutiva guarda con la risa.

(i) Propondré una caracterización de un cierto tipo de humor que denominaré “verbal” como herramienta epistémica. Para ello, echaré un vistazo rápido a las principales teorías del humor, adaptando para ello la clásica distinción entre teorías de la superioridad, la liberación y la incongruencia (Larkin-Galiñanes-2017) y deteniendo en un puñado de autores localizables en esta última: Eliot-1856, para quien el chiste nos permite descubrir relaciones entre ideas que en un principio parecen completamente desconectadas, la caracterización psicológica del proceso de comprensión del chiste de Jerry-1972 como un problema cognitivo que exige una reordenación de nuestras expectativas en orden a eliminar la incongruencia, y especialmente la comparación de Ludwig Wittgenstein entre la investigación filosófica y la comprensión de un chiste (tal y como es interpretada en Vilanova-2013 o Garmendia 2021).

(ii) Exploraré la naturaleza de la emoción asociada al humor. Partiendo del dato curioso de que tanto en la psicología popular como en la científica no hay ningún concepto específico del tipo “emoción de humor”, más allá de lo que vagamente pueda vincularse con la risa (en el círculo de las emociones de Russell-1980 lo más cercano sería “happy” o quizás “relaxed), y de que apenas hay literatura sobre ello. La causa de esta ausencia bien podría ser el difícil encaje dentro de las concepciones más tradicionales de la emoción: simplificando, una explicación que reduce las emociones a percepciones o sensaciones fisiológicas (como la llamada teoría de James-Lange, James-1884) deja fuera el logro cognitivo asociado a la comprensión del texto, una explicación que reduzca emociones a juicios o evaluaciones (como Kenny-1963 o Nussbaum-2001) no acaba de apresar la vivencia subjetiva que acompaña a la risa. Más prometedoras a este respecto son concepciones híbridas como la teoría perceptiva de las emociones de Prinz-2004 (donde la percepción directa de un estado fisiológico es la percepción indirecta de un objeto o estado del entorno causal o representacionalmente ligado al primero) o Tappolet- 2016 (donde la emoción es directamente la percepción evaluativa de un objeto), pero aquí se hace muy difícil encajar el aspecto intelectual de la captación del chiste, ya que en ambos casos se

* vilanova@filos.ucm.es

adscribe al tipo de percepción propia de la emoción una naturaleza no conceptual (o al menos no directamente conceptual, como en el caso de Tappolet).

(iii) Extraeré del repaso previo un diagnóstico provisional: cualquier teoría o tesis sobre el humor es en mayor o menor grado parcial respecto al fenómeno de estudio: o deja fuera la “intelección” asociada a la comprensión del chiste, o deja fuera la “emoción” asociada a la risa, o no deja clara la relación entre ambas. Propondré que en buena medida las dificultades se derivan de dos modelos deficientes que funcionan en el trasfondo de la mayoría de enfoques teóricos:

- un modelo clásico de conocimiento en la que las emociones no entran en juego salvo como impedimentos o trabas para el correcto discurrir del proceso epistémico,
- un modelo clásico de las emociones como contenidos o estados de la conciencia (condicionada por la tendencia a la cosificación que impone el modelo “nombre-objeto” del significado), que choca con la variedad de componentes (psicológico, fisiológico, evaluativo, conductual, expresivo, cognitivo...) que aparecen en su examen empírico.

Frente a estos modelos, intentaré fundamentar la tesis de que la emoción forma parte constitutiva del proceso epistémico, sirviendo a la vez como “criterio” y “manifestación” del conocimiento. Me basaré fundamentalmente en las observaciones que hace al respecto Wittgenstein en sus consideraciones sobre seguir una regla, aunque también haré mención a otras concepciones epistémicas “amables” con las emociones como las guías emocionales de Morton 2010.

Complementariamente, defenderé la tesis de que la emoción es una entidad compleja, en la que hay una continuidad entre aspectos fenoménicos, cognitivos, fisiológicos, conceptuales y físico-naturales, y que no puede adscribirse solo al individuo sino al todo de la situación, incluyendo en ella el trasfondo natural, histórico, cultural y hasta institucional. Para ello me basaré de nuevo en las consideraciones que hace al respecto de las emociones Wittgenstein en las Investigaciones y en los escritos dedicados a la filosofía de la psicología, aunque también prestaré atención a la teoría enactivista de las emociones (Colombetti-2014).

(iv) Estaré ya en condiciones de presentar mi tesis. En ella jugará un papel protagonista la distinción popular entre “entender (el texto de) el chiste” y “pillar el chiste”, siendo el criterio de este último la aparición de la emoción de la que es a la vez parte y manifestación la risa. Consideraré que el humor (al menos el humor verbal) opera (sin salirse de ellas) con las propias reglas de nuestro lenguaje y las prácticas asociadas, produciendo interpretaciones inusitadas y aplicaciones inesperadas (tan legítimas como las estándares desde un punto de vista lógico), sacando de esta manera a la luz posibilidades y conflictos latentes. Por ello, un buen chiste requiere y a la vez propicia un conocimiento sofisticado y profundo de las reglas de juego (del mismo tipo del involucrado en la investigación filosófica). Ahora bien, para “pillar el chiste” no es suficiente con entender como

su conclusión surge de las propias reglas, sino que es necesario apreciar el carácter intrínsecamente absurdo, insensato o “inútil” de la misma (lo cual, a su vez, tiene a la vez el efecto terapéutico de ayudarnos a aceptar la naturaleza limitada y conflictiva de nuestros conceptos y el efecto pedagógico de entrenarnos en la excepcionalidad y la flexibilidad de pensamiento). Por ello la comprensión solo se produce cuando aparece la emoción asociada a la percepción del absurdo y con ella o incluyéndola a ella la risa.

Finalmente, resumiré una conclusión en una tesis: la risa constituye la percepción sofisticada de aspectos profundos de nuestras reglas lingüísticas. Tesis que aclararé explorando distintos sentidos en que se puede entender el “constituye”: como “es constitutiva de”, o “está interpenetrada con”, o “es indesligable de”, o “es a la vez criterio, manifestación y parte pública de”.

Bibliografía

- [1] Larkin-Galiñanes, Cristina (2015). “An Overview of Humor Theory” en Salvatore Attardo (ed.), *Routledge Handbook of Language and Humour*, Routledge.
- [2] Barrett, L. F., & Russell, J. A., Eds. (2015). *The psychological construction of emotion*, Guilford Press.
- [3] Berger, Peter (1997). *Redeeming Laughter: The Comic Dimension of Human Experience*, De Gruyter.
- [4] Colombetti, Giovanna (2014). *The Feeling Body: Affective Science Meets the Enactive Mind*, MIT Press.
- [5] Döveling, Katrin, von Scheve, Christian, Konijn, Elly A.(2010). *The Routledge Handbook of Emotions and Mass Media*, Routledge.
- [6] George Eliot (1856). *German Wit*, Westminster Review, vol. LXV: 1-33.
- [7] Fehr, Beverley and Russell, James A. (1984). “Concept of Emotion Viewed from a Prototype Perspective”, *Journal of Experimental Psychology: General*, 113(3): 464-486.
- [8] James, William (1884). “What is an Emotion?”, *Mind*, 9(2): 188-205.
- [9] Jerry Suls, Paul (1972). “A two-stage model for the appreciation of jokes and cartoons: an information process analysis”, en Goldstein y McGhee (eds.), *The Psychology of Humour*, Pergamon Press.
- [10] Kenny, Anthony (1963). *Action, Emotion and Will*, Routledge and Kegan Paul.
- [11] Lewis, Michael y Haviland-Jones, Jeannette y Feldman Barrett M. Lisa (2016, cuarta edición). *Handbook of Emotions*, The Guilford Press.

- [12] Morton, Adam (2010). “Epistemic emotions”, en Goldie 2010: 385-400.
- [13] Nussbaum, Martha C. (2001). *Upheavals of Thought: The Intelligence of Emotions*, Cambridge University Press.
- [14] Peirce, C. S. (1877). “The Fixation of Belief”, *Popular Science Monthly* 12, pp. 1-15) Prinz, Jesse (2004): *Gut Reactions: a Perceptual Theory of Emotion*, Oxford University Press.
- [15] Peter Goldie (ed.) (2010). *The Oxford Handbook of Philosophy of Emotion*, Oxford: Oxford University Press.
- [16] Provine Robert R (2016). “Laughter as a scientific problem: An adventure in sidewalk neuroscience”, *Journal of Comparative Neurology* , 524(8): 1532-9.
- [17] Robert R. Provine (2001). *Laughter: an scientific investigation*. Penguin.
- [18] Russell, James A. (1980). “A Circumplex Model of Affect”, *Journal of Personality and Social Psychology*, 39(6): 1161–1178.
- [19] Santiago Garmendia (2021). “Wittgenstein sobre chistes y filosofía”, en Lastres y Bassols (eds.): *Wittgenstein y la ciencia* (en prensa).
- [20] Tappolet, Christine (2016). *Emotions, Values, and Agency*, Oxford University Press.
- [21] Vilanova, Javier (2013). “La profundidad filosófica de un chiste”, *Apeiron*, 6: 90.105.

Relativamente gracioso

Antonio Blanco Salgueiro*
Universidad Complutense de Madrid

En esta contribución se explorarán dos temas: la universalidad *versus* diversidad (o relatividad) de humor y la posible relación de dependencia del humor con respecto al lenguaje. En relación con lo primero, no se pondrá en duda que el humor es un fenómeno humano universal, y que se ha avanzado bastante en el proyecto de formular una teoría plausible acerca de sus orígenes evolutivos y de su función en nuestras mentes y en nuestras vidas (para esto me apoyaré básicamente en el ya clásico Hurley, Dennett & Adams, 2011). Sin embargo, se insistirá en que las explicaciones universalistas dejan mucho margen para explorar las distintas fuentes de diversificación en el humor humano. Con respecto a la relación entre humor y lenguaje, frente a teorías como la de Hurley, Dennett & Adams (2011) que no dan mucho protagonismo al lenguaje (sino más bien a las emociones), se defenderá, en primer lugar, que sin la capacidad del lenguaje (no específica de lenguas particulares) de recombinar elementos libremente no sería posible o apenas sería posible el humor. Con ello, se defiende en el ámbito del humor la idea de que las características que parecen exclusivas de los humanos se derivan de que somos la única especie con lenguaje (cf. Bickerton, 1995). En segundo lugar, se conectará el tema de la relatividad del humor con distintas formas de la clásica y controvertida hipótesis de la relatividad lingüística (cf. Blanco Salgueiro, 2017). La aplicación a campos nuevos de esta hipótesis es una tendencia filosófica en auge en los últimos tiempos (cf. Blanco Salgueiro, 2021). Además de la relatividad del humor debida a la diversidad entre las lenguas (que se muestra del modo más evidente en la intraducibilidad de muchos chistes), examinaremos otros tipos de diversidad lingüística (y también cultural e individual) que pueden intervenir en la diversificación del humor. Completaremos nuestro examen con un esbozo de lo que podría ser una teoría de los infortunios para el humor, inspirada en la teoría de los infortunios que Austin ofrece para los realizativos. Austin sitúa el *bromear* como un caso paradigmático de *etiolation* (decoloración) del lenguaje. El objetivo es estudiar las condiciones en las que esta decoloración del lenguaje puede funcionar y aquellas en las que fracasa de maneras diversas, de modo que el humor podría presentarse como relativo a ciertas condiciones de éxito que serían la cara positiva de las formas en las que puede fallar (y falla).

* ablancos@filos.ucm.es

Bibliografía

- [1] Austin, J. L. (1962). *How to Do Things with Words*. Oxford: Oxford University Press.
- [2] Bickerton, D. (1995). *Language and Human Behavior*. Seattle: University of Washington Press.
- [3] Blanco Salgueiro, A. (2017). *La relatividad lingüística (variaciones filosóficas)*. Madrid: Akal.
- [4] Blanco Salgueiro, A. (2021). “The Many Facets of Linguistic Relativity”. En P. Stalmaszczyk (ed.): *The Cambridge Handbook of the Philosophy of Language*. Cambridge: Cambridge University Press, pp. 675-691.
- [5] Hurley, M. M., Dennett, D. y Adams, R. B. (2011). *Inside Jokes. Using Humor to Reverse-Engineer the Mind*. Cambridge, Massachusetts / Londres: MIT Press.

La contestación de las dicotomías: la sonrisa de Austin

Saleta de Salvador Agra*
Universidad Complutense de Madrid

La posición de Austin contra las dicotomías (Austin, 1964), su sospecha de hipersimplificación y apriorismo (Sbisà, 1973), será el punto de partida para pensar el supuesto binarismo de los usos “serios” y “no serios” del lenguaje. Mi propuesta se centrará entonces en sumar a la discusión de las dicotomías, entre “verdadero/falso” y “valorativo/ factual” (Austin, 1962), aquella que el propio filósofo enuncia en la *Lecture II* al hilo de la discusión sobre las condiciones de felicidad de los performativos. De tal manera que, sobre el trasfondo de la famosa contienda respecto a la interpretación austiniana capitaneada por Derrida y Searle, esbozaré tres posibles líneas de continuación del debate tomando el caso del humor como pretexto. Así:

i) Por una parte, expondré aquellas claves que podrían conducirnos a pensar en un “nuevo comienzo” de *How to Do Things with Words*. Esto es, añadir a los dos comienzos (en orden a las denominadas “teoría especial” y “teoría general” a las que remite el recorrido lineal de las conferencias; del inicio hasta la *Lecture VIII* y desde ésta hasta el final) la posibilidad de uno nuevo que relea la obra desde la no binariedad de lo serio/no serio. Tomando como guía las preguntas del por qué usa el humor y qué hace Austin con él, examinaré los innumerables ejemplos que, con claro tono jocoso, abundan en su obra (el matrimonio con un mono, el bautizo de pingüinos o perros, la promesa de una zanahoria a un burro o el caso de matar a un gato con manteca, entre otros), así como la presumida falta de seriedad en sus ingeniosos títulos (como en el ya mencionado *How to do Things...* o “A plea for excuses”, etc.) además de introducir la noción de “humor performativo” para volver a leer *How to Do Things with Words* como un acto de habla en sí mismo. Es decir, no sólo reparando en lo que dice sobre los actos de habla sino entendiéndolo como uno más.

ii) Por otra parte, consideraré tres figuras humorísticas con las que se ha identificado a Austin (Eíron, Don Juan y Charles Chaplin), para indagar en lo que, considero, podría apuntar a diferentes modos del “salirse con la suya” (Austin, 1962). En consecuencia, me detendré, en primer lugar, en el eíron (Rodríguez, 2019) de la comedia griega antigua, quien con ironía, empleando palabras con su sentido antónimo, al fingir consigue sembrar dudas; en segundo lugar, en el Don Juan de Moliere (Felman, 1980) como el seductor que

* saletade@ucm.es

hace promesas para no cumplirlas o mantenerlas y, por último, en el del icónico Charles Chaplin (Zwargerman, 2010) como ejemplo de humor negro y estratégico que nos empuja a la necesidad de contemplar las cosas desde otra perspectiva.

iii) Finalmente, analizaré el caso del “humor performativo” que se da cuando el juego del lenguaje resultante tiene que ver, precisamente, con difuminar las fronteras y los límites que establecen las dicotomías. Esto me permitirá estudiar la seriedad del humor de la mano de algunas propuestas feministas donde, por ejemplo, las imágenes de las mujeres humoristas consiguen revertir jerarquías y deshacer aquellas dicotomías que tanto disgustaban a Austin (Berlin, 1984). Al tiempo que, en línea con la propuesta de relectura de la obra austiniana, repararé en lo apropiado del uso del humor para una investigación filosófica “seria” en general (Vilanova, 2015) y feminista en particular. De tal forma que la potencia activa del acto de habla del humor, como un “hacer reír” más que un “decir”, nos conducirá a analizar el reclamo de Butler, cuando dejó escrito que “la risa ante categorías serias es indispensable para el feminismo” (Butler, 1990).

En conclusión, la tesis que aquí trataré de defender pretende situarse en línea con “la contestación de las dicotomías” (Sbisà, 1973), viendo así como la sonrisa de Austin dibuja justamente el camino que nos obliga a regresar a la raíz de la falsa dicotomía entre los usos serios y los no serios del lenguaje.

Referencias bibliográficas

- [1] Austin, John (1962). *How to do Things with words*, Oxford University Press, Oxford.
- [2] Austin, John (1964). *Sense and Sensibilia*, Oxford University Press, Oxford.
- [3] Berlin, Isaiah (1980). “J. L. Austin and the Early Beginnings of Oxford Philosophy” in *Personal Impressions*, The Hogarth Press, London.
- [4] Butler, Judith (1990). *Gender Trouble: Feminism and the Subversion of Identity*, Routledge, New York.
- [5] Derrida, Jacques (1988). *Limited Inc*, Northwestern University Press, Evanston.
- [6] Felman, Shoshana (1980). *Le Scandale du corps parlant. Don Juan avec Austin ou La séduction en deux langues*, Seui, París.
- [7] Navarro Reyes, Jesús (2010). *Cómo hacer filosofía con palabras. A propósito del desencuentro entre Searle y Derrida*, Fondo de Cultura Económica: Madrid.
- [8] Rodríguez Ponce, M^a Isabel (2019). “El eíron y otras reminiscencias griegas en la filosofía de John L. Austin”, *Synthesis*, vol. 26, n^o2, pp. 1-17

- [9] Sbisà, Marina (1973). “Una visione del mondo per J. L. Austin” in Giacon, C. (ed.) *Filosofia y Politica e altri saggi*, Editrice Antenore, Padova.
- [10] Vilanova, Javier (2015). “La profundidad filosófica de un chiste”, *Apeirón*, nº2, pp. 90- 105.
- [11] Zwargerman, Sean (2010). *Wit’s End: Women’s Humor As Rhetorical and Performative Strategy*, University of Pittsburgh Press, Pittsburgh.

Siempre nos quedará la risa

Antonio Duarte Calvo*
Universidad Complutense de Madrid

Alejándonos del concepto “inteligente” del *humor* y de la dificultad teórica que se genera al tratar de delimitar este término (ver, por ejemplo, Attardo 1994: 1-13), en esta contribución exploraremos la “estupidez abductiva” como detonador de *lo cómico* o, simplemente, de *la risa*. En este sentido, seguiremos la distinción de Hurley, Dennet y Adams (2011: 117), donde, básicamente, el humor se asocia con algún tipo de circunstancia semántica donde existe un error y uno logra descubrirlo, mientras que la risa se produce simplemente por el placer de encontrar un error particular en nuestros sistemas de creencias. Por tanto, en los juegos del lenguaje, por ejemplo, en la habilidad para entender un chiste en relación con el complejo mundo que nos rodea, ¿qué ocurre si el intérprete supone mal o elige una hipótesis irrelevante? ¿Qué ocurre, en definitiva, cuando la interpretación falla?

El simple hecho de “no pillar un chiste” nos remite a múltiples factores que pueden estar fallando en el proceso de comunicación: la capacidad de “predicción” del hablante, la interpretativa del oyente o la falta de una teoría previa adecuada de los interlocutores (Davidson 2005, Duarte 2019). Además, atendiendo a la dimensión social del suceso, este tipo de errores en el proceso de interpretación donde existe un trasfondo humorístico puede ser fácilmente utilizado como una herramienta para la promover el escarnio y la exclusión social (Bell 2017).

En esta comunicación nos centraremos en la rigidez interpretativa, ayudándonos de la teoría del humor mecanizado de Bergson. Las agudas consideraciones en torno a lo cómico de Bergson en *La risa* (Bergson, 2003) nos llevarán a visualizar la “estupidez abductiva” que lleva, por un lado, a una interpretación fallida para el humor y, por otro, a una situación cómica. Para Bergson, lo que mueve a la risa, por ejemplo, cuando un hombre tropieza y cae, es la rigidez mecánica frente a la agilidad despierta y flexibilidad viva del ser humano. Estas dos fuerzas complementarias, tensión y elasticidad, son las que se ponen en juego en la comedia. A partir del cuerpo torpe o mecanizado o de la pobreza de espíritu, vendrán todas las variedades del envaramiento y la pobreza psicológica. Así una expresión rígida en el rostro es la que nos hace pensar en algo ridículo frente a la movilidad que caracteriza a la fisonomía. Por eso, Bergson afirma que «las actitudes, gestos y movimientos del cuerpo humano son risibles en la exacta medida en que este cuerpo

* antduart@ucm.es

nos hace pensar en un simple mecanismo» (Bergson, 2003: 31). Es algo que claramente explotan las caricaturas y los humoristas, en general. La vida, podríamos decir, frente al automatismo instalado. Por eso la imitación hace reír, porque de algún modo, se rescata la parte automática del imitado.

Estas cuestiones son transferibles al acto comunicativo. La interpretación mecánica, rígida de una proferencia que reclama flexibilidad y viveza, es decir, el automatismo en la interpretación provoca la risa. El “olfato de lo que es relevante” es esencial para interpretar el discurso humorístico basado, principalmente, en las desviaciones intencionales del lenguaje. En este sentido, cuando se pierde este olfato, cuando se pierde la capacidad de llegar a la adecuada reconstrucción abductiva, se transmuta lo vivo por lo mecánico y se produce lo cómico. Como afirma Wirth, «la “abducción estúpida” ofrece una interpretación que es sorprendentemente irrelevante, incoherente, y, o muy compleja o muy simple, y por tanto, estúpida» (Wirth, 2000: 6).

El propio Bergson cuenta una anécdota que nos puede servir para ejemplificar estas cuestiones: alguien comentó delante de Boufflers, oficial y poeta francés del siglo XVI-II, aludiendo a un presuntuoso personaje, “corre en pos del ingenio”. A esto, parece que Boufflers respondió: “Apuesto por el ingenio”. Si nos parece ingeniosa la respuesta, nos hace sonreír, al menos, es porque toma la metáfora “correr en pos del ingenio” de una forma literal, como si el tipo presuntuoso y el ingenio estuvieran compitiendo a ver si se alcanzan. Aquí Boufflers juega hábilmente con la metáfora reclamando de nosotros una interpretación viva y flexible. Otra cosa hubiera sido, y hubiera llevado a risa, que alguien, haciendo gala de un pobre proceso abductivo/interpretativo, hubiera preguntado, “¿pero, cómo va a apostar por el ingenio? ¡Eso es imposible!”.

En definitiva, en relación con la interpretación lingüística, “lo cómico” surge al verse defraudadas nuestras expectativas, siendo, en estos casos, el opuesto de “lo humorístico”. Dos caras de una misma moneda, que, en cualquier caso, nos llevan a la risa.

Referencias

- [1] Attardo, S. (1994). *Linguistic Theories of Humor*. Mouton de Gruyter, Berlin / Nueva York.
- [2] Bell, N. D. (2017). “Failed humor”. En S. Attardo (ed.), *The Routledge Handbook of Language and Humor*, Routledge, Nueva York / Londres.
- [3] Bergson, H. (2003)/[1899]. *La risa. Ensayo sobre la significación de lo cómico*. Losada, Buenos Aires.
- [4] Davidson, D. (2005)/[1986]. “A Nice Derangement of Epitaphs”. *Truth, Language and History*, (pp. 89-108), Clarendon Press, Oxford.

- [5] Duarte, A. (2019). “On abduction and interpretation”. *Crítica, Revista Hispanoamericana de Filosofía*, Vol. 51, No. 151 (abril 2019), pp. 65–84.
- [6] Hurley, M. M., D. Dennett, y R. B. Adams. (2011). *Inside Jokes. Using Humor to Reverse-Engineer the Mind*. Cambridge, Massachusetts / Londres: MIT Press.
- [7] Wirth, U. (2000). “Abduction, Wit, Stupidity – from Peirce to Freud”. In M. Bergman & J. Queiroz (Eds.), *The Commens Encyclopedia: The Digital Encyclopedia of Peirce Studies. New Edition*. Pub. 121219-1823a.

SYMPOSIUM:
THE CONTEXT-SENSITIVITY
OF EXPRESSIVES

General description and justification of interest

21st century philosophy of language has witnessed an increasing interest in expressive meaning, that is, in those aspects of natural language that are primarily used to express one's attitudes, rather than convey factual information. Expressions with derogatory content have occupied a central role in the discussion. This symposium is devoted to a special type of derogatory expressions, namely pejoratives, such as 'jerk' and 'asshole'. The literature on those, particularly Potts (2005, 2007), Schlenker (2007), and Tonhauser, Beaver, Roberts and Simons (2013), converges on the idea that for an expressive to be felicitously uttered in a context, it is enough that the speaker should have a derogatory attitude toward the target of the expressive, regardless of whether this attitude is shared by the interlocutors, or whether the target actually deserves to be held in low esteem. Recently, Cepollaro, Domaneschi and Stojanovic (2020) have presented two experimental studies that seriously challenge this idea. In a nutshell, the results show that expressives are highly sensitive to contextual information. What is more, this holds not only for expressives in their referential use (that is, constructions of the form "that jerk X") but also in their predicative use (that is, "X is a jerk"). The presentations at this symposium aim to engage with the findings of these studies. They discuss the implications of the findings for the existing theories of expressives, and address further issues raised by pejoratives, regarding the nature of the derogatory content associated with them, the differences between referential and predicative uses, and their place in a broader space of negatively valenced expressions.

Cited works

- [1] Cepollaro, Bianca, Domaneschi, Filippo, and Stojanovic, Isidora (2020). When Is It OK to Call Someone a Jerk? An experimental investigation of expressives, *Synthese*. <https://doi.org/10.1007/s11229-020-02633-z>
- [2] Potts, Christopher (2005), *The logic of conventional implicatures*. Oxford: OUP.
- [3] Potts, Christopher (2007), The expressive dimension, *Theoretical Linguistics*, 33, 165-197.
- [4] Schlenker, Philippe (2007), Expressive presuppositions, *Theoretical Linguistics*, 33 (2), 237-245.
- [5] Tonhauser, Judith, Beaver, David, Roberts, Craige, and Simons, Mandy (2013), Toward a taxonomy of projective content, *Language*, 89 (1), 66-109.

Literally “a jerk”. An experimental study of the semantics of negative expressives

Bianca Cepollaro*
Università Vita-Salute San Raffaele

Filippo Domaneschi**
Department of Philosophy - University of Genoa

Isidora Stojanovic***
Institut Jean-Nicod, CNRS

Background

The philosophical semantic literature on expressive terms such as “bastard” and “jerk” converges on two claims (Macià 2002; Potts 2005, 2007; Schlenker 2005). The first one is the idea that the use of such terms does not place any special constraints on the conversational context (Tonhauser et al. 2013). The second claim is that the content associated with the expressives is *speaker-oriented*. The expressive content of a sentence like (1a) or (1b) amounts to something like (1c):

- (1) a. $\llbracket \text{That jerk } x \rrbracket^c$
b. $\llbracket x \text{ is a jerk} \rrbracket^c$
c. λx : the speaker in c feels negatively about the target x .

In our study, which focuses on predicative – rather than referential – uses of “stronzo” (tr. “jerk”), we challenge both claims.

Experiment

Participants. 219 participants took part in the study [MA= 31.01; SD=12.51; 156f; 63m].

Stimuli. We created 8 written vignettes in Italian composed of a *context scenario* and a *target sentence* of the form “ x is P ” where x is the target individual and P is the negative expressive “stronzo”. We also included 18 filler items: 9 positive (e.g., “kind”, “wise”)

* bianca.cepollaro@gmail.com

** filippo.domaneschi@unige.it

*** isidora.stojanovic.nicod@gmail.com

and 9 negative (e.g., “impolite”, “lazy”) non-expressive fillers. The context sentence was manipulated in order to generate 5 conditions – see Fig. 1. In the subjective condition (SBJ), it was part of the story that the speaker calling another character “a jerk” had a negative attitude towards her/him. In the first intersubjective condition (Int SBJ 1), both the speaker and the other interlocutors had a negative attitude towards the target, while in the second intersubjective condition (Int SBJ 2) such a negative attitude was entertained by other unspecified persons. The objective condition (OBJ) did not include information about how the characters felt about the target, but it was part of the story that the target is an underhanded and deceitful person. Finally, we included a neutral condition (NEU) providing no contextual information supporting the use of the pejorative, i.e. no information as to whether the target did something bad or people feel negatively about her/him. Filler items (both positive and negative fillers) were presented in neutral conditions only, i.e. in non-supporting contexts.

Procedure. Participants had to read the context scenarios and were then asked to rate on a 5-point Likert scale (1 unacceptable; 5 completely acceptable) the degree of acceptability of the target sentence as uttered by a fictional speaker and of the positive and negative fillers in the 5 experimental conditions. The rationale was to explore which contextual condition made the use of “stronzo” more acceptable, i.e., which condition best supported the utterance containing the negative expressive.

Results. Figure 2 shows the mean acceptability rate for each of the five experimental conditions. A non-parametric Kruskal-Wallis test was conducted to examine the differences on participants’ acceptability judgement of the target sentences according to experimental condition. This analysis revealed a significant effect of condition ($\chi^2(4) = 180$; $p < 0.001$). Pairwise comparisons between conditions were conducted using the Dwass-Steel-critchlow-Fligner (DSCF) method. Most importantly, this analysis revealed that participants judged as significantly more acceptable the target sentences containing the target expression when these were presented in condition OBJ than in conditions Int_Sbj_1 ($W = 13.83$; $p < 0.001$), Int_Sbj_2 ($W = 12.29$; $p < 0.001$), NEU ($W = 18.44$; $p < 0.001$), and SBJ ($W = -12.07$; $p < 0.001$). Additionally, the target sentences were judged as significantly less acceptable when these were presented in condition NEU than in conditions Int_Sbj_1 ($W = -7.38$; $p < 0.001$), Int_Sbj_2 ($W = -6.35$; $p < 0.001$), and SBJ ($W = 4.61$; $p = 0.01$) – see Table 1 for the results of all pairwise comparisons. A second set of statistical analyses was conducted to explore potential acceptability differences between participants’ judgement of the target sentences in the neutral condition and the positive and negative filler sentences – see Figure 3 for the mean acceptability rates in the three conditions. A Kruskal-Wallis test revealed a significant effect of condition ($\chi^2(2) = 134$; $p < 0.001$). DSCF pairwise comparisons revealed that filler sentences were significantly more acceptable in the positive than in the negative condition ($W = 9.72$; $p < 0.001$). Most importantly, participants judged as less acceptable the target sentences in the neutral condition as compared to both the positive ($W = 16.05$; $p < 0.001$) and the negative ($W = -7.17$; $p < 0.001$) filler items.

Conclusion

Our study challenges both of the claims shared by most theories of expressives: it challenges the first claim in that it shows that expressives in their predicative use are sensitive to contextual information, and that they are judged less acceptable than other negatively valenced, evaluative terms, such as “unbearable”. Furthermore, it challenges the second claim in that the use of expressives was deemed more acceptable in the objective condition (“X is an underhanded and deceitful person”) rather than in the subjective or intersubjective conditions (“The speaker dislikes X”, “The speaker and others dislike X”): this speaks against the idea that the content associated with expressives is speaker-oriented “The speaker feels negatively about the target”.

References

- [1] Macià, Josep (2002), Presuposición y Significado Expresivo, *Theoria: Revista de Teoria, Historia y Fundamentos de la Ciencia*, 3 (45), 499-513.
- [2] Potts, Christopher (2005), *The Logic of Conventional Implicatures*. Oxford: OUP.
- [3] Potts, Christopher (2007), The Expressive Dimension, *Theoretical Linguistics*, 33, 165-197.
- [4] Schlenker, Philippe (2007), Expressive Presuppositions, *Theoretical Linguistics*, 33 (2), 237-245.
- [5] Tonhauser, Judith, Beaver, David, Roberts, Craige and Simons, Mandy (2013), Toward a Taxonomy of Projective Content, *Language*, 89 (1), 66-109.

| CONTEXT SENTENCE | | | TARGET SENTENCE |
|---|-------------|---|--------------------------------------|
| Sara, Luca and Chiara are at the pub with some friends Marco is drinking a coffee in another table | SBJ | Luca hates Marco | Luca says to Sara: Marco is [a] jerk |
| | INTER-SBJ 1 | Sara, Luca and Chiara hate Marco | |
| | INTER-SBJ 2 | Sara, Luca, Chiara and other persons hate Marco | |
| | OBJ | Marco must have done something bad | |
| | NEU | // | |

Figure 1. Example of a target item with the context sentences generating the five experimental conditions

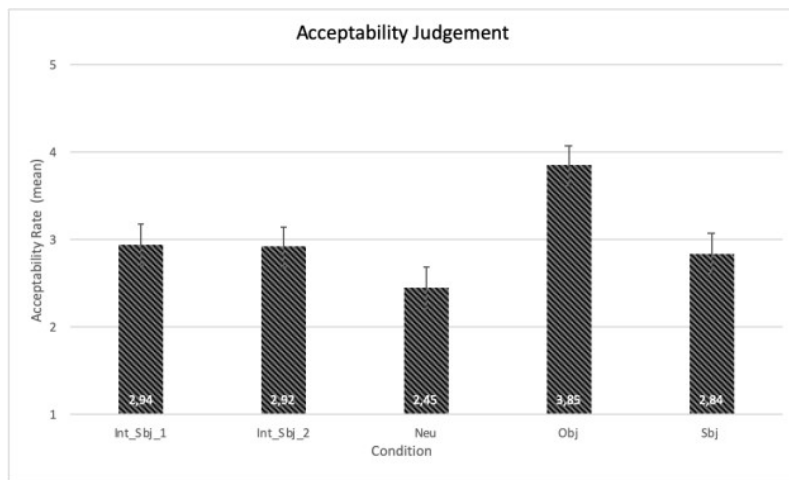


Figure 2. Mean acceptability rate for each of the five experimental conditions

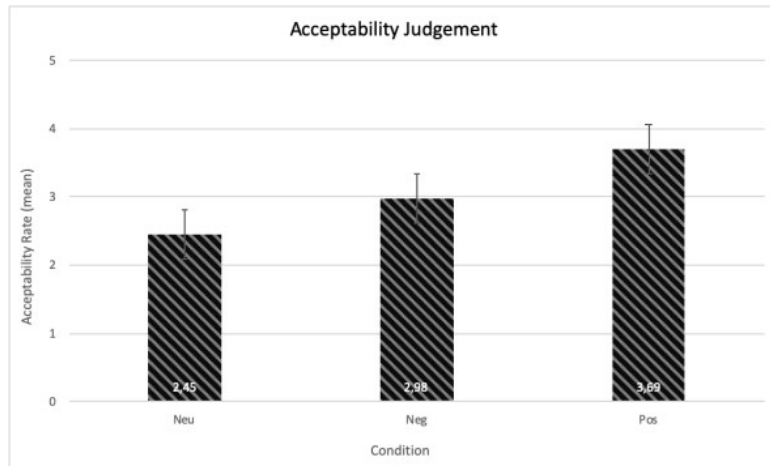


Figure 3. Mean acceptability rate of the target item in neutral condition (NEU), negative fillers and positive fillers

The pragmatics of all-purpose pejoratives

Víctor Carranza*
Ecole Normale Superieure

This presentation argues that all-purpose pejoratives such as ‘jerk’ or ‘bastard’ (also called ‘epithets’) are just plain vanilla descriptions of personality traits that are generally seen as impairing for the self and for interpersonal relationships across different contexts. It is standardly assumed that all-purpose pejoratives are expressive (Kaplan 1998, Potts 2005, Cepollaro et al. 2020). Expressives are a class of terms whose main function is to display some kind of emotion, mostly of the speaker. Typical examples of expressives are thus expletive adjectives (1a) and all-purpose pejoratives (1b):

- 1) a. The **fucking** dog is on the couch.
- b. That **bastard** Kaplan was promoted.

Expressives display a set of specific properties. First, the information conveyed by expressives is independent, that is, irrelevant to the descriptive content of the utterances in which they occur. As (2a-b) illustrates, expletive adjectives and all-purpose pejoratives can be omitted without altering the sentence’s truth-conditions. Second, the information conveyed by expressives is projective. As (2a-b) also illustrates, an inference about the speaker’s feelings is triggered even though the relevant item occurs under the syntactic scope of an if-clause:

- 2) a. If the **fucking** dog is on the couch, John will be upset.
 = If the dog is on the couch, John will be upset. (*truth-conditional vacuity*)
 →The speaker feels negatively about the dog. (*projective inference*)
- b. If that **bastard** Kaplan was promoted, then the regents acted foolishly.
 = If Kaplan was promoted, then the regents acted foolishly.
 →The speaker feels negatively about Kaplan.

However, it has been observed that this characterization is problematic in the case of all-purpose pejoratives (Beller 2013). In atomic sentences, all-purpose pejoratives convey information that contributes to the utterance’s truth-conditions and that remains in the scope of truth-conditional operators (e.g. if-clauses). In (3), ‘bastard’ does not seem to trigger an inference about the speaker’s emotions:

* rcarranza@pucp.pe

- 3) If Kaplan is a **bastard**, then he will be promoted.

It may be suggested that all-purpose pejoratives are ambiguous between expressive and non-expressive readings according to whether they occur as modifiers in definite descriptions or as simple predicates. However, appealing to syntactic ambiguities doesn't seem promising. Even though pejorative modifiers are assumed to manifest an exclusive preference for non-restrictive (i.e. truth-conditionally irrelevant) interpretations, they can give rise to restrictive interpretations in certain contexts. In particular, restrictive readings arise in contexts where the pejorative modifier directly addresses a QUD (4a), or else when its content has an additional explanatory function (4b) (Martin 2014, Esipova 2020):

- 4) a. Anne: Which one of Alex's ex-boyfriends did you see?
 Kyle: The **bastard** one.
- b. Alex eliminated all her **asshole** contacts from Facebook.
 → Alex eliminated those contacts *because* they were assholes.

In view of the problems that emerge in expressive accounts, Beller (2013) argues that all-purpose pejoratives are better categorized as evaluative predicates such as 'fun' or 'tasty'. Evidence that all-purpose pejoratives are evaluative is that they give rise to 'faultless' disagreements. As the exchange in (5) illustrates, in discussions about whether someone is a bastard or not neither participant has necessarily made a mistake (is at fault). Therefore, Beller (2013) proposes to capture the evaluative character of all-purpose pejoratives such as 'bastard' by incorporating a judge parameter on their denotation.

- 5) Anne: John is a **bastard**.
 Kyle: No, he is not.

A problem with this approach assumes that all-purpose pejoratives such as 'asshole' don't encode anything descriptive. That is, that all-purpose pejoratives are 'descriptively ineffable' (Potts 2005) and therefore that categorizing someone as an asshole just amounts to saying something about how one feels about such individual. This assumption is incorrect. In a behavioral study, Hyatt et al. (2019) investigated the psychological traits associated with people categorized as assholes, dicks and bitches. It emerged from this study that participants shared a conception about the kind of individual these insults refer to. Prototypical assholes, dicks and bitches are those that are arrogant, distrustful, selfish and manipulative. Therefore, relativisation to a judge parameter in the case of all-purpose insults go against evidence that these are descriptive of observable personality traits, and not just indicative of the speaker's emotions.

But then, if all purpose pejoratives are descriptive, how evaluative inferences arise? All-purpose pejoratives refer to personality traits that are commonly believed to be undesirable and impairing for the self and for interpersonal relationships. Thus, uses of terms such

as ‘jerk’ or ‘bastard’ will be typically linked to negative evaluations. In other terms, the evaluative content associated to all-purpose pejoratives arises as a ‘generalized’ conversational implicature (Levinson 2000): it is common knowledge that the personality traits referred to by all-purpose pejoratives are evaluated negatively and that utterances that include such terms reflect those negative evaluations, so saying that someone is a jerk or bastard will typically express the speaker’s own emotions about the individuals targeted.

It might be argued that this pragmatic approach doesn’t capture the distinctive kind of derogation inherent to the use of all-purpose pejoratives, which typically serve as an outlet for the speaker’s most immediate emotions and not just to indicate that some individual is distrustful, selfish, etc. However, in our pragmatic theory such expressive ‘flavor’ can be considered a consequence of the highly colloquial status of these terms. The idea is that a token of ‘jerk’ not only refers to distrustful or selfish behaviors, but also constitutes a breach of politeness standards. This transgression itself indicates that the intensity of the negative evaluation is higher than normal. This may explain why, as mentioned above, all-purpose pejoratives have been standardly classified alongside profanities such as ‘fucking’ or ‘damn’ in the literature, despite their enormous linguistic differences: both constitute a breach of politeness standards that pragmatically implicates that the speaker intends to express an intense emotion.

Bibliography

- [1] Beller, C. (2013). Manufactured and inherent pejorativity. *Semantics and Linguistic Theory*, 23:136.
- [2] Cepollaro, B., Domaneschi, F. & Stojanovic, I. (2020). When is it ok to call someone a jerk? An experimental investigation of expressives. *Synthese*.
- [3] Esipova, M. (2020). *Composure and composition*. MS.
- [4] Hyatt, C., Maples-Keller, J. L., Sleep, C., Lynam, D., and Miller, J. (2017). *The anatomy of an insult: Popular derogatory terms connote important individual differences in externalizing behavior*.
- [5] Kaplan, D. (1999). The meaning of *ouch* and *oops*: Explorations in the theory of meaning as use, 2004 version. MS, University of California, Los Angeles.
- [6] Levinson, S. C. (2000). *Presumptive Meanings: The Theory of Generalized Conversational Implicature*, Cambridge, MA: MIT Press.
- [7] Martin, F. (2014). Restrictive vs. nonrestrictive modification and evaluative predicates. *Lingua*, 149:34-54.
- [8] Potts, C. (2005). *The logic of conventional implicatures*. Oxford Studies in Theoretical Linguistics Oxford: Oxford University Press.

Jeremy is a... Expressive-Relativism and Expressive Predicates in Copular Positions

Justina Berskyte*
University of Manchester

Abstract

Broadly construed, expressives are words that convey a speaker's attitude towards an entity or a situation. Following Potts (2005, 2007), expressives are said to have the property of independence - removal of an expressive from a proposition does not affect its truth-conditional content. This idea has been challenged by authors who point to examples where the expressive seems to contribute directly to the truth-conditional content of the proposition (see for example, (Geurts, 2007), (Gutzmann, 2015), (McCready, 2010)). To illustrate this consider the following:

Attributive: That fuckhead Jeremy forgot the turkey.

Copular: Jeremy is a fuckhead.

In both examples, the predicate *fuckhead* communicates some expressive content - the negative attitude of the speaker. However, only in Copular does it appear to contribute to the truth-conditional content. Removing the predicate from Copular will affect the truth-conditional content, the proposition will be left incomplete. The task is to explain the semantics of the word *fuckhead* when it seemingly behaves wildly differently in different syntactic positions.

In the first part of this talk I'll consider Potts' claim that "[a]ll predicates that appear in copular position must necessarily fail to be expressive" (Potts, 2007, 194). I will quickly dismiss Potts' claim and show that treating Copular as purely descriptive is unsatisfactory. Firstly, it doesn't explain any relationship between *fuckhead* in Copular and Attributive. Secondly, and more importantly, it fails to account for the highly important expressive element of *fuckhead* in Copular.

I'll go on to consider a more promising account put forth by Gutzmann (2015, 2016). Gutzmann combines Expressivist semantics with Contextualism - I call this view Expressive-Contextualism. By adopting a hybrid semantics Gutzmann is able to account for both expressive and descriptive dimensions of cases like Copular. Although Gutzmann's theory has a lot more explanatory power than that of a pure Expressivist account, I raise

* justina.berskyte@manchester.ac.uk

concerns over adopting a Contextualist semantics as the correct semantics for expressive predicates.

Lastly, I will present a novel hybrid account that combines Expressivist semantics with Relativism (Lasersohn, 2017). I call this view *Expressive-Relativism*. The Relativist semantics will allow us to account for the truth-conditions of sentences like Copular. We'll say that the utterance is true just in case the judge (most typically the speaker) thinks that Jeremy is a *fuckhead*. The expressive content – the speaker's negative attitude – will be captured via Expressivist semantics. I demonstrate that Expressive-Relativism can explain the relationship of *fuckhead* in Attributive and Copular. It can also provide a plausible story for the truth-conditional aspect of *fuckhead*.

References

- [1] Geurts B (2007) Really fucking brilliant. *Theoretical Linguistics* 33(2):209-214.
- [2] Gutzmann D. (2015) *Use-Conditional Meaning: Studies in Multidimensional Semantics*. Oxford University Press, Oxford: UK.
- [3] Gutzmann D (2016) If Expressivism Is Fun, Go for It. In: Meier C, van Wijnberger-Huitink J (eds) *Subjective Meaning: Alternatives to Relativism*, De Gruyter, Berlin, pp. 21-46.
- [4] Lasersohn P (2017). *Subjectivity and Perspective in Truth-theoretic Semantics*. Oxford University Press, Oxford.
- [5] McCready E (2010). Varieties of Conventional Implicature. *Semantics and Pragmatics* 3(8): 1-57.
- [6] Potts C (2005). *The Logic of Conventional Implicatures*. Oxford Studies in Theoretical Linguistics, Oxford University Press, Oxford.
- [7] Potts C (2007) The Expressive Dimension. *Theoretical Linguistics* 33(2): 165-198.

Thick Ethical Concepts – The Extent and Possible Explanations of the Polarity Effect

Lucien Baumgarten*
University of Zurich

Kevin Reuter**
University of Zurich

Pascale Willemsen***
University of Zurich

Philosophers distinguish between two types of evaluative concepts: *thin* and *thick*. *Thin* evaluative concepts, such as ‘good’ or ‘bad’, evaluate an entity positively or negatively, yet alone they provide no information as to why the entity is evaluated in this way. *Thick* concepts provide such descriptive information. For instance, ‘rude’ not only conveys a negative evaluation of a person or their behavior, but a person’s behavior can only count as rude if it fulfills various descriptive criteria, such as hurting someone’s feelings, causing offense, or violating norms of proper conduct.

In a recent paper, Willemsen and Reuter investigate how the evaluation of a thick concept is communicated – whether it is by semantic or pragmatic means. Their results suggest that neither the semantic nor the pragmatic view can fully account for the use of thick terms in ordinary language. More strikingly though, Willemsen and Reuter detect an asymmetry between positive and negative thick concepts, such that denying the positive evaluation communicated by a positive thick concept is less contradictory than denying a negative evaluation of a negative thick concept. Such a polarity effect had not been predicted by any of the theoretical accounts of thick concepts and raises deep philosophical questions about whether we should even think of thick concepts as one homogenous class.

In this talk, we examine the extent of the polarity effect as well as potential explanations of it. First, extending the experimental paradigm as Willemsen and Reuter (2021), we demonstrate that the polarity effect occurs in different embeddings, including single entity attributions, limited scope attributions, as well as generic statements (Exp. 1). Here is an example for ‘rude’ and its opposite ‘friendly’ for all three embedding conditions:

- Single Entity Attribution: “*Amy is rude [/virtuous], but by that I am not saying something negative [/positive] about her.*”

* lucien.baumgartner@philos.uzh.ch

** Kevin.Reuter@philo.unibe.ch

*** pascale.willemsen@uzh.ch

- Limited Scope Attribution: *“Some people are rude [/friendly], but by that I am not saying something negative [/friendly] about them.”*
- Generic Statement: *“People are rude [/friendly], but by that I am not saying something negative [/positive] about them.”*

After reading one of these statements, participants answered the question of whether the speaker contradicts herself. We tested six pairs of terms and found a significant polarity effect when comparing positive and negative terms, but also when comparing each individual positive term with its negative counterpart.

How can we explain this effect? It might be suspected that there is something about the cancellability test used in Willemsen and Reuter (2021) and Experiment 1 that triggers different interpretations of sentences containing positive and negative thick terms. For instance, it seems plausible to think that a positive statement in which the positive evaluation is denied triggers a pro tanto interpretation of goodness and badness. When the speaker says that Amy is friendly, but denies that this is something positive, she might want to communicate that just because Amy is friendly, that doesn't make her a good person overall. The same reinterpretation might not be triggered in the same way by a negative term. Being rude is usually enough to make an overall negative statement about a person as a whole. To engage with these and related worries, we applied different methodological tools. In Experiment 2, we used a different experimental setting in which we gave participants statements of the following form:

Michelle has the task of assessing Sandra's behavior. She writes in her report: “What Sandra did was rude but bad [friendly but good].”

Participants then rated how odd or natural this statement sounded to them. Instead of cancelling the evaluation usually conveyed by the thick concept, we create an explicit contrast between a thick concept and the evaluation usually expressed. Participants found it significantly more natural when a contrast between positive thick concepts and goodness was created (friendly but good), compared to a contrast between negative thick concepts and badness (rude but bad). This indicates, in line with Willemsen and Reuter (2021) and Experiment 1, that the positive evaluation of a positive thick concepts is less strongly communicated than the negative evaluation of a negative thick concept.

Alternatively, one might think that Amy is considered to be friendly to an extent that makes her friendliness problematic for her – too much friendliness might make her an easy victim of rude people who take advantage of her. Such an understanding of the target sentence is less readily available for negative terms. It seems that no degree of rudeness would ever make this a positive trait of a person. We already believe that this explanation is made less likely by Experiment 2, but we cannot rule it out entirely. Thus, we explored the viability of this hypothesis by modifying Experiment 1. In Experiment 3, we use the following stimulus which makes it even more explicit that the speaker wants to make a non-evaluative statement:

“What Tom did last week was rude [friendly], but by that I am not saying something positive or negative about his behavior that day. I mean this in a fully neutral way.”

Our results demonstrate again a significant difference between positive and negative thick concepts. For a negative thick concept, canceling the evaluation and using the term purely neutrally sounds more contradictory in comparison to positive thick concepts.

Based on our results, we discuss the possibility that what drives the polarity effect are differences in the acceptability of using positive and negative evaluative language – even to merely describe a person or their behavior.

ISBN: 978-84-09-35510-5



9 788409 355105