

## PERSONAS EN EL MUNDO: LA PERSPECTIVA DE PRIMERA PERSONA Y EL NATURALISMO

Agustín Vicente & Adrián Sampedro León<sup>1</sup>

### RESUMEN

En el presente artículo examinamos diversas respuestas a la cuestión de qué somos, fijándonos sobre todo en las propuestas reduccionistas y eliminativistas sobre las personas o los *yoes*. Concluimos que, a día de hoy, parece más razonable el dualismo que el naturalismo, si por tal cosa entendemos la reducción o eliminación de entidades psicológicas.

**PALABRAS CLAVE:** Persona, yo, naturalismo, dualismo, relación de constitución, experiencia y su unidad.

### ABSTRACT

In this paper we examine different answers to the question of what we are, focusing in particular on eliminative and reductivist proposals about persons or selves. We conclude that, as of today, dualism seems more reasonable than naturalism, if by naturalism we understand the thesis that psychological entities can be reduced or eliminated.

**KEYWORDS:** Person, self, naturalism, dualism, the constitution relation, experience and its unity.

### 1. INTRODUCCIÓN

---

<sup>1</sup>Este trabajo es fruto de la colaboración. El orden en que aparecen sus autores es aleatorio.

Cuando decimos cosas como “[Yo] creo que [yo] estoy cogiendo miedo a volar”, usamos el pronombre de primera persona para referirnos a algo, pero ¿a qué nos referimos? Hay tres respuestas que tienen cierta pujanza. La primera es que nos referimos a una persona; la segunda es que nos referimos a un animal humano (o, alternativamente, a un cuerpo o a un cerebro; en definitiva, a algo material); y la tercera es que no conseguimos referirnos a nada realmente existente. Las dos primeras respuestas cuentan con el favor del sentido común, aunque una sea no naturalista y la otra sea naturalista. La tercera respuesta hace considerablemente más daño a los oídos, pero parece tener cada vez más predicamento en ciertos círculos naturalistas. En lo que sigue, expondremos estas tres posturas y los argumentos más importantes a favor de cada una de ellas. Nuestro propósito es mostrar que no es sencillo deshacerse de (la intuición de que existen) las personas, al tiempo que es difícil ubicar a éstas en el mundo natural.

## 2. ¿PERSONAS O ANIMALES?

John Locke fue quien primero distinguió entre personas y hombres o mujeres -lo que a día de hoy llamamos ‘animales humanos’ (Locke, 1690, Libro segundo, XXVII) Para él un hombre o una mujer es una entidad biológica con determinadas características funcionales. Podemos decir que un hombre o una mujer es un tipo específico de organización funcional. Las personas, sin embargo, son otra cosa. Las personas son seres racionales que pueden pensar en sí mismos como ellos mismos a lo largo del tiempo: una persona es “a thinking intelligent being, that has reason and reflection, and can consider itself as itself, the same thinking thing, in different times and places” (Locke, 1690, Libro segundo, XXVII; 9). Un rasgo definitorio de las personas, por tanto, es la autoconciencia o pensamiento de orden superior, la capacidad de pensar acerca de sus estados mentales en tanto estados mentales *proprios*.

Según las versiones recientes más depuradas (Baker, 1998; Baker 2000) la marca distintiva de las personas es la capacidad de exhibir una perspectiva “fuerte” de primera persona, por oposición a la perspectiva “débil”. La perspectiva débil o “rudimentaria” de primera persona es aquella en virtud de la cual ciertas criaturas (paradigmáticamente organismos sintientes de suficiente complejidad neurofisiológica, como mamíferos no humanos) se representan la realidad a través de experiencias y estados mentales básicos de actitud proposicional. Bajo esta perspectiva egocéntrica, el organismo es el centro

receptor de tales experiencias y estados mentales, que permiten explicar su conducta mediante silogismos prácticos. La perspectiva fuerte o “robusta” de primera persona es la capacidad de concebirse a sí mismo como propietario exclusivo de la perspectiva débil, y de conceptualizar la distinción entre la primera y la tercera persona. Esta distinción se refleja a nivel lingüístico. En oraciones simples como ‘[Yo] estoy cogiendo miedo a volar’ el pronombre ‘Yo’ comporta meramente la referencia de primera persona. Sin embargo, en oraciones compuestas como “[Yo] creo que [yo] estoy cogiendo miedo a volar”, el verbo psicológico asociado al sujeto precede a una cláusula subordinada con la misma estructura gramatical, lo que comporta, además de la referencia de primera persona, la *atribución* de primera persona. Es esta atribución de primera persona la que revela la facultad de auto-concebirse como sujeto de estados mentales frente al resto del mundo. Cabe añadir que sólo cuando hay una segunda aparición del pronombre de primera persona la expresión ‘yo’ refiere, en ambas ocurrencias, a una persona (Baker, 2000).

Cabe preguntarse si caracterizar a dos entidades (animales humanos y personas, en nuestro caso) de diferente manera es suficiente para sostener que se trata de cosas diferentes. Efectivamente, en sí no es suficiente. Pero de la distinta caracterización se derivan condiciones de persistencia distintas, y eso sí que parece ser un buen motivo para trazar una distinción ontológica. Un hombre o una mujer siguen siendo el mismo/la misma mientras mantengan el mismo tipo de organización funcional. En particular, esto quiere decir que “están ahí” desde los primeros estadios del feto hasta que mueren. La persona no “está ahí” hasta que el animal humano tiene cierta edad, y también puede desaparecer antes de que desaparezca el hombre o la mujer. Por otra parte, la persona puede sobrevivir a la desaparición del animal humano: si el cerebro del animal se trasplanta a otro, o si de algún modo se consigue que un robot reciba, retenga y ponga en uso la información contenida en el cerebro del animal, la persona puede perfectamente pervivir, mientras que el animal humano podría desaparecer.

La idea principal de los lockeanos contemporáneos, como Baker, Lowe (1996), y otros, es que personas y animales humanos son dos tipos de cosas diferentes, ambos existentes. Los conceptos de persona y de animal humano son ambos conceptos *sortales* (tipificadores). Lo característico de los conceptos o clasificaciones sortales es que proporcionan condiciones de identidad y de persistencia de los objetos que caen bajo ellos: te dicen cuándo un determinado objeto es el mismo que otro que quizás puedas creer diferente, así como

cuándo un determinado objeto deja de existir. Por otra parte, según muchos filósofos (Wiggins, 1980), los conceptos sortales sirven para caracterizar las esencias de los objetos. Para cada objeto, particular o individuo del mundo, hay al menos, y a lo sumo, un concepto sortal que lo individúa esencialmente. Por ejemplo, aunque podemos categorizar una hoja de un árbol de muchas maneras, sólo su clasificación como hoja nos proporciona su caracterización esencial: la esencia de una hoja de un árbol es ser una hoja (y no, por ejemplo, brillar con la luz de primavera, estar pendiente de una rama, o ni siquiera ser una entidad viva). Las propiedades capturadas por estos conceptos de substancia o sortales pueden denominarse ‘propiedades de clase primaria’ (Baker, 1999).

De modo que el lockeano típicamente mantiene dos tesis: (a) que los conceptos de persona y de animal humano son conceptos sortales diferentes, i.e., que las personas y los animales humanos tienen distintas propiedades de clase primaria y (b) que lo que nos caracteriza a nosotros esencialmente es ser personas, no ser animales humanos. La primera tesis se apoya en que las condiciones de identidad y persistencia que nos proporcionan el sortal de persona son diferentes a las condiciones de identidad y persistencia que nos proporcionan el sortal de animal humano (o, alternativamente, de cuerpo humano): como hemos visto, la persona, *qua* persona, puede sobrevivir a la desaparición del animal o del cuerpo, mientras que el animal, *qua* animal, también puede sobrevivir a la desaparición de la persona (pensemos en un individuo en estado vegetativo), y puede existir sin necesidad de que exista la persona (cuando el animal está en estado de feto). Lo mismo ocurre, por otra parte, si en lugar de contraponer personas y animales, contraponemos personas y cerebros, preguntándonos si acaso las personas no son sus cerebros: el cerebro puede morir y la persona pervivir (si, por ejemplo, logramos mantenerla anclada a otro soporte). De hecho, nuestro concepto de persona parece permitirnos especular con la idea de que las personas en tanto tales sobrevivan incluso en ausencia de soporte físico: quizás estaríamos dispuestos a considerar a Dios como una persona, y no nos cuesta nada pensar en los fantasmas en términos personales –un fantasma, básicamente, es una persona que ha “perdido” su cuerpo-. No estamos diciendo que estas entidades ficticias sean personas –no lo son porque no existen-; sólo indicamos que el concepto de persona es tal que podríamos aplicarlo a entidades incorpóreas si éstas existieran.

La segunda tesis es que cuando, como en el ejemplo de '[yo] creo que [yo] estoy cogiendo miedo a volar', utilizamos el pronombre de primera persona, nos referimos no al animal humano, sino a la persona. La razón para mantener tal cosa es que somos personas, no animales humanos. Como se ha dicho, hay dos tipos de clasificaciones sortales que se nos aplican en principio: la clasificación "animal humano" y la clasificación "persona". Si los conceptos sortales, como dicen muchos filósofos, nos proporcionan la esencia de los objetos, particulares o individuos, entonces tenemos que optar por una de estas dos clasificaciones como la clasificación que nos caracteriza. La segunda tesis afirma que somos esencialmente personas, esto es, seres racionales que pueden pensar en sí mismos como ellos mismos. Como hemos visto, consideramos que las personas tienen condiciones de identidad y persistencia diferentes a las de sus cuerpos (o a las de los animales humanos). La cuestión entonces es: ¿con qué nos identificamos a nosotros mismos, con el animal que muere o la persona que lo sobrevive en el caso, por ejemplo, de ser "trasplantada"? ¿Con la persona que desaparece en el estado vegetativo o con el animal que la sobrevive? Parece que la respuesta en ambos casos es que nos identificamos con la persona, no con el animal.

El filósofo que con mayor fuerza ha defendido la postura "animalista" en los últimos tiempos es Eric Olson. Sus principales argumentos hacen referencia a lo problemático que puede resultar multiplicar las entidades. Por ejemplo, pongamos el caso sencillo de quién pesa qué (Olson, 1998): digamos que tú pesas 70 kilogramos, y que tú, como dicen los "personalistas" eres una persona. Cohabita contigo un animal humano que también pesa 70 kilogramos. ¿Quiere esto decir que entre los dos pesáis 140? Obviamente no, pero ¿entonces? Entonces resulta que no hay dos cosas, sino sólo una. El mismo problema se suscita con respecto a las propiedades mentales. Digamos que tú piensas que la mejor forma de viajar es en tren. Ese 'tú' se refiere a la persona que eres. Pero resulta que el animal que cohabita contigo también piensa: tiene un cerebro que le permite hacerlo. ¿Estáis entonces los dos pensando lo mismo? ¿Cada uno piensa dos ejemplares del mismo pensamiento? ¿Sólo la persona piensa? ¿Sólo el animal lo hace? Multiplicar las entidades implica afrontar complicaciones innecesarias, que podemos evitar simplemente absteniéndonos de añadir al animal humano, que sabemos positivamente que existe, otra entidad sobre la que podemos albergar dudas.

Como hace notar el propio Olson (2007), sin embargo, este tipo de problema ha de afrontarlo cualquier teórico, y no sólo el personalista. El

problema, en el caso del animalista, es el siguiente: junto contigo, el animal humano, cohabita un conjunto de materia (carne y huesos, por así decir). Tú pesas 70 kilogramos, y esa materia también. Tú piensas que el mejor modo de viajar es el tren, y esa materia también. ¿Entonces? Si el personalista ha de renunciar a las personas a partir de este argumento, parece que el animalista debería renunciar al animal, y quedarse con la materia que constituye al animal. Es decir, el animalista, al final, no está en mejor situación que el personalista.

### 3. ¿PERSONAS CONSTITUIDAS POR CUERPOS?

La propuesta más destacada para bloquear el problema de la duplicación de propiedades es la teoría de la constitución material (Baker, 2000). Según Baker, el problema estriba en explicar cómo los criterios de identidad y las condiciones de existencia de los organismos humanos y las personas difieren, al tiempo que se reconoce que las personas se hallan íntimamente unidas por una relación de dependencia con los organismos humanos. La respuesta provista por Baker es que la relación entre personas y cuerpos no es ni de identidad ni de existencia separada, sino de *constitución*.

La idea nuclear de la constitución, según Baker, es ésta: cuando ciertas entidades con ciertas propiedades de clase primaria se hallan en determinadas circunstancias favorables, nuevas entidades, con nuevas propiedades de clase primaria (con nuevas condiciones de persistencia) cobran existencia (ver detalles técnicos y presentación formal en Baker, 2000, p. 43; y la última versión en Baker, 2003, p. 35). Así, un determinado trozo de mármol, cuando se inserta en un mundo como el nuestro, se convierte en el constituyente de una nueva entidad, la estatua, cuya propiedad primaria extrínseca es ser una obra de arte. En lo concerniente a las personas humanas, las circunstancias favorables incluyen las condiciones intrínsecas (cierto grado de complejidad neurofisiológica) y las condiciones extrínsecas (ciertas condiciones ambientales propicias) conducentes al desarrollo y preservación de la perspectiva fuerte de primera persona. En principio, dado que la persona resultante coincide espaciotemporalmente con el organismo humano que la constituye, parece que la teoría de la constitución no es capaz de responder a la objeción de la duplicación de propiedades. Sin embargo, la doctrina de la constitución material permite compatibilizar la convicción preanalítica de que existen dos entidades numéricamente distintas (aunque espaciotemporalmente solapadas) con la repulsa de la consecuencia señalada por Olson. Esto es así porque la co-

ubicación espaciotemporal de dos entidades numéricamente distintas es congruente con la exigencia fisicista de que la entidad constituyente agote materialmente a la entidad constituida (Wiggins, 1968). Baker es quien ha proporcionado la explicación más detallada de la relación entre las propiedades de las entidades constituyente y constituida mediante las nociones de posesión independiente, posesión derivada y préstamo o herencia de propiedades, que a continuación presentamos.

*Posesión independiente:* Una entidad  $x$  posee una propiedad  $F$  independientemente de sus relaciones de constitución con otra entidad  $y$  en un instante de tiempo  $t$  si y solamente si  $x$  instancia  $F$  en  $t$  y o bien (i)  $x$  constituye  $y$  en  $t$  y el hecho de que  $x$  instancie  $F$  no implica que  $x$  constituya entidad alguna en  $t$ , o bien (ii)  $y$  es constituido por  $x$  en  $t$ , y el hecho de que  $y$  instancie  $F$  no implica que  $y$  esté constituido por una entidad que pueda instanciar  $F$  sin constituir algo en  $t$ .

*Posesión derivada:* Una entidad  $x$  posee una propiedad  $F$  derivadamente en  $t$  si y solamente si (i) existe un  $y$  tal que  $y$  ejemplifica  $F$  independientemente de sus relaciones de constitución con  $x$  en  $t$  y (ii)  $x$  no ejemplifica  $F$  independientemente de sus relaciones de constitución con  $y$  en  $t$ .

Para ilustrar las definiciones con ejemplos: una persona ejemplifica la propiedad de ser portadora de responsabilidades penales independientemente de sus relaciones de constitución con su organismo, ya que la propiedad de tener responsabilidades penales no podría ser ejemplificada por un organismo que no constituyera una persona. Un organismo, por su parte, es portador de responsabilidades penales derivadamente, i.e., únicamente en virtud de sus relaciones de constitución con una entidad que ejemplifica tal propiedad independientemente. En contraste, un organismo pesa 70 kilogramos independientemente de sus relaciones de constitución, ya que la propiedad de pesar 70 kilogramos es compatible con el hecho de que el organismo que ejemplifica tal propiedad no constituya entidad alguna (una persona). Una persona pesa 70 kilogramos derivadamente, i.e., únicamente en virtud de sus relaciones de constitución con un organismo que pesa 70 kilogramos independientemente.

*Herencia de propiedades:* Cuando la instanciación de una propiedad F por parte de una entidad x es dependiente de sus relaciones de constitución con otra entidad y, se dice que x hereda la propiedad F, instanciándola derivadamente.

Así, parece que la doctrina de la constitución permite armonizar la intuición de que las personas *son* sustancias materiales con la elusión de la consecuencia de la proliferación de propiedades indicada por Olson. En efecto, no hay duplicación de propiedades (en este caso propiedades físicas como ‘pesar 70 kilogramos’), porque las personas heredan las propiedades físicas de los organismos que las constituyen (su instanciación de tales propiedades es derivada o dependiente de su relación de constitución con los organismos). La acusación de Olson reposa en una lectura errónea de la cópula ‘es’ como ‘es’ de *predicación* en lugar del ‘es’ de *constitución*. Sólo la primera lectura entraña la duplicación de propiedades.

De la concepción constitucionalista de Baker parece desprenderse un dualismo no cartesiano, en que las personas son (‘es’ de predicación) entidades psicológicas –en cuanto portadoras de perspectiva fuerte de primera persona- y son (‘es’ de constitución) entidades materiales –en cuanto constituidas por organismos físicos-. Sin embargo, esta virtud lleva aparejado un problema, al que dedicaremos el resto de la sección.

La cuestión es la siguiente: si asumimos el préstamo bidireccional de propiedades, parece que hemos de aceptarla consecuencia de que las personas son compuestas y divisibles en virtud de sus relaciones de constitución con sus cuerpos humanos. Ahora bien si, esta consecuencia choca con la asunción de los dos siguientes principios, independientemente razonables:

(PSE) *Principio de singularización de experiencias* (Strawson, 1959): la carga posesiva involucrada en las condiciones de singularización o individuación de las experiencias presupone analíticamente un sujeto como propietario de tales experiencias.

(PUE) *Principio de la unidad de la experiencia* (Lowe, 2009): el sujeto de la experiencia referido en PSE es unitario, y no hay entidad que pueda figurar como sujeto unitario de un conjunto de experiencias si al menos algunas de estas experiencias pueden existir en su ausencia.



Estos dos principios hacen difícil sostener que las personas, tal y como son concebidas por Baker, puedan ser las portadoras unitarias de estados mentales, dado que, por el principio de herencia de propiedades, parece que habría que decir que las personas tienen partes; a saber, las partes que componen los organismos que las constituyen. Baker (2000, 2002) mantiene que ni los organismos ni las personas constituidas por ellos son los portadores exclusivos de todas las experiencias y estados mentales: “On the constitution view, the mind is not a substance at all; the relevant substances are person and human organism, both of which are fully material, and neither of which is the exclusive bearer of mental properties” (2002, p. 33). Esta afirmación de Baker se funda en la presunción de que no todas las propiedades mentales son causadas en exclusiva ni por las personas ni por los cuerpos que las constituyen. Ambos –personas y cuerpos– generan ciertos grupos de propiedades mentales específicamente asociados con ellos y, consiguientemente, ambos son portadores de propiedades mentales. No obstante, cabe reafirmarse en que es propio de las experiencias conscientes depender conjuntamente de un sujeto que es propietario unificador de tales experiencias. Una cosa es que los estados mentales dependan causalmente del organismo, y otra que esos mismos estados mentales dependan metafísicamente de un sujeto (que es, al fin y al cabo, lo que los convierte en subjetivos) que es el mismo para todos ellos. Ambas tesis son compatibles, y ambas, en realidad, parecen ser abrazadas por Baker, quien, como hemos dicho al inicio, concibe la perspectiva fuerte o “robusta” de primera persona como la capacidad de concebirse a sí mismo como *propietario exclusivo* de la perspectiva débil. Ahora bien el problema es que, como ni el cuerpo humano como un todo ni cualesquiera de sus partes pueden constar como únicos poseedores unitarios de experiencia, tampoco puede constar como sujeto unitario de la experiencia una entidad que está hecha (por herencia) de las mismas partes de que está hecho el cuerpo humano.

Comencemos con la idea de que ni el cuerpo ni ninguna de sus partes puede ser sujeto unitario de la experiencia. La idea consiste en mostrar que al menos algunos estados mentales pueden subsistir a la desaparición del organismo como un todo o de cualquiera de sus partes (recuérdese el principio PUE). Por ejemplo, las alucinaciones generadas por los “miembros fantasma” son pruebas en favor de que los estados mentales pueden pervivir a la desaparición de partes del organismo (y por tanto del organismo como un todo). Sin embargo, puede defenderse la misma tesis sin la apelación a tales

situaciones extremas: si a un individuo se le amputase un dedo de la mano, se vería privado de ciertos *qualia* asociados a ese dedo, pero continuarían existiendo la inmensa mayoría de estados mentales, aunque el organismo como un todo haya desaparecido. Si apostamos por la tesis más débil según la cual los estados mentales dependen unitariamente, si no del organismo como un todo, del cerebro como un todo, la estrategia anterior puede aplicarse de nuevo con el mismo resultado. Finalmente, tampoco parece plausible sostener que todo estado mental depende de una parte determinada del cerebro (de una determinada área cerebral). Es decir, la totalidad de la experiencia no depende ni de una parte del organismo ni del organismo como un todo.

La única objeción razonable que parece puede dirigirse a Lowe es que su argumento descansa sobre la dudosa presuposición según la cual la sustracción de una parte de mi cuerpo implica que mi cuerpo, como un todo, deja de existir. No obstante, hay medios para salvar esta objeción sin aprobar el esencialismo mereológico. El argumento de Lowe discurre así (Lowe, 2009): si se admite que no todos mis estados mentales (llámense 'E') dependen de mi cuerpo con todas sus partes, al que se denominará 'C', entonces todos los estados mentales que no dependen de C, dependerán del mismo cuerpo *menos* una de sus partes. Llámese a este cuerpo 'D'. Supóngase que se identifica mi persona con el cuerpo D, que es el sujeto de todos mis estados mentales E. Pregúntese entonces sobre qué fundamento puede D ser considerado sujeto de E, preferentemente a C, dado que E no depende de la parte que C incluye y de la que D carece. ¿No es la diferencia física entre C y D simplemente irrelevante para la proclamación de uno de ellos como sujeto de E? Sin embargo, en tal caso, se ha de responder ante una disyuntiva: o bien tanto C como D son legítimos sujetos de E, o bien ninguno de ellos lo es. No podemos intentar solventar el problema por la primera vía, ya que C y D son entidades numéricamente distintas, y los estados mentales E dependen de un solo sujeto-yo mismo-. Se ha de concluir, por tanto, que ni C ni D son los sujetos de E y, así, que yo soy el sujeto de E y, como tal, distinto de C y D. Mis estados mentales dependen de mí como sujeto unitario.

La conclusión de Lowe es que mis estados mentales dependen holísticamente, no de mi cuerpo o cualesquiera partes integrantes del mismo, sino de mí como poseedor unitario de experiencias, puesto que sólo la desaparición del yo como sujeto de pensamiento comporta la desaparición de todos y cada uno de mis estados mentales. Un corolario es la concepción atómica de las personas: todo portador genuino de estados mentales ha de ser

*simple*, lo cual parece, en primera instancia, entrar en conflicto con el dualismo constitucionalista de Baker, ya que éste implica que la persona está, aunque derivadamente, conformada por partes y, por tanto, no puede figurar como entidad psicológica de la que dependen cabalmente todas sus experiencias.

Un posicionamiento realista sobre estados mentales comporta, pues, dos compromisos. El primero (PSE) es la aceptación de *un sujeto* de experiencia, dado que éste forma parte esencial de las condiciones de individuación de estados mentales. El segundo, que se deriva de (PUE), es la exigencia de que tal sujeto de experiencia sea *indivisible*, dado que la dependencia de los estados mentales con respecto a su portador es *unitaria*, y no parece haber objeto físico compuesto que resulte ser aspirante legítimo a portador del que dependen unitariamente todas sus experiencias. Estos dos compromisos parecen dar al traste con el mecanismo de herencia de propiedades propuesto por Baker, y, por tanto, hacen resurgir el problema de la duplicación de propiedades.

A nuestro juicio, en la confrontación con el “animalismo”, la postura de autores como Lowe y Baker lleva las de ganar (al fin y al cabo, como hace notar el propio Olson, 2007, el problema de la duplicación de propiedades también afecta al animalista). No deja de ser curioso, por otra parte, pues casi desde la aparición de la filosofía de la mente como disciplina distinguida, se ha asumido que el dualismo de sustancias es falso, girando la discusión en torno a si las propiedades mentales son o no propiedades físicas/naturales. La particular “visión heredada” nos ha dicho que todos los instanciadores de propiedades son o bien entidades físicas básicas, o bien compuestos de ellas, mientras que la duda es si las propiedades mentales son igualmente explicables en términos de entidades físicas y relaciones no problemáticas como la de composición (Kim, 1998). Quizás descubramos ahora que el dualismo de sustancias no es tan disparatado como creíamos (para defensas recientes del dualismo de sustancias frente al de sólo propiedades, Zimmerman, 2010 y Lycan, 2013). Y de hecho, el dualismo de sustancias no cartesiano que defiende Baker no parece disparatado en absoluto.

De modo que el filósofo naturalista tiene que buscarnos otra ubicación. No puede simplemente identificarnos con animales humanos, ni con cuerpos ni con cerebros, y mucho menos con sumas de moléculas. ¿Existe alguna otra solución? Aparentemente sí: puede sostener que ese pronombre personal del que hemos hablado no refiere a una persona, sino a una ficción.

#### 4. ¿PERSONAS O ILUSIONES?

Existe, *prima facie*, una respuesta a los argumentos presentados en las secciones anteriores, aparejada a una respuesta a la cuestión principal de qué somos. La idea básica es que todo esto de lo que hemos hablado no es sino una ilusión generada por nuestro cerebro. El cerebro construye un yo, al que le atribuye, entre otras cosas, el ser sujeto unitario de la experiencia. Hacia los dos/tres años de edad, comienza este proceso de construcción, que arranca con el autorreconocimiento (e.g. reconocerse ante un espejo) y el uso de los pronombres, y crece hasta concebir el yo no sólo como el sujeto de la experiencia y el agente de las acciones, sino como una entidad con historia y rasgos característicos (sociales, morales, etc.). El yo, entonces, adquiere dos dimensiones: una vinculada al aquí y ahora, que percibe nuestras percepciones, realiza nuestras acciones, piensa nuestros pensamientos, etc.; y otra relacionada con nuestra autoconcepción. Sin embargo, para muchos, todo este yo robusto no es sino una ilusión adaptativa generada por nuestro cerebro (Robins et al., 2008)

¿Por qué es adaptativa? En primer lugar, es funcionalmente adaptativa, o, por ponerlo de otra manera, es funcionalmente exitosa. Tomemos el caso de la agencia: un sistema tiene que distinguir los cambios originados por él mismo de los cambios que ocurren en el mundo. Por ejemplo, tiene que distinguir un cambio en lo que percibe ocasionado por el movimiento de sus ojos de uno causado por un súbito movimiento del mundo en el que focalizaba su mirada. Lo mismo ocurre con las acciones: tanto el mundo que percibimos como nuestro propio cuerpo cambia cuando emprendemos una acción. Sin embargo, el mundo que percibimos y nuestro propio cuerpo pueden sufrir cambios independientemente de lo que hagamos. Hemos de distinguir el primer tipo de cambios del segundo. Una buena estrategia es situar el origen de los cambios imputables al sistema en el sistema como un todo, y proporcionar esa información al propio sistema. De ahí surge que las acciones se sientan como propias, como llevadas a cabo por uno mismo. Es más útil –y quizás sea necesario incluso- atribuir todas las acciones a un todo que imputar la responsabilidad a cada subsistema particular. La información está así disponible para todos los subsistemas, y es más sencillo seguir la pista de las acciones llevadas a cabo (y compararlas, rectificarlas, planificar, etc.) e integrarlas con el resto de conocimiento que tenemos sobre nuestra interacción con el mundo en general.

Crear en un yo nos sirve también para el propósito de auto-regularnos (o de que el sistema se auto-regule). Tenemos planes de todo tipo: algunos a muy corto plazo y otros que pueden llevarnos toda una vida. Los planes entran en conflicto constantemente, y tenemos que decidir, por ejemplo, si damos prioridad al plan inmediato, o lo cancelamos teniendo a la vista los planes a más largo plazo. Para hacer esto es útil creer en un yo que perdura al que le atañen –e incluso configuran– todos los planes en juego. La idea de que somos entes esencialmente psicológicos que albergan una perspectiva sobre el mundo y tienen creencias, intenciones, deseos, planes, etc., también nos sirve en la tarea de interpretar el comportamiento de organismos similares a nosotros. Entendernos como personas a nosotros mismos y a los demás facilita la comprensión de los otros. Es, seguramente, la manera más efectiva de poder vivir en comunidad (Robins, et al., 2008). Es más, para poder vivir en comunidad no sólo necesitamos poder entender a los otros que forman parte de esa comunidad, sino que también necesitamos poder responsabilizarles de sus acciones. Los sistemas morales, según muchos, han evolucionado para mantener las comunidades cohesionadas. No puede haber un sistema moral, sin embargo, si no hay individuos –personas– a los que imputar las acciones. La responsabilidad no recae sobre cerebros o partes de cerebros, sino sobre lo que consideramos sus poseedores, que tienen condiciones de persistencia moral, por así decirlo, más robustas.

Sin embargo, no es suficiente tener una explicación de cómo una ficción podría ser útil para sostener que, efectivamente, se trata de una ficción. Lo dicho anteriormente es compatible con que el yo no sea una ficción, sino algo muy real. El ficcionalista tiene que explicar, primero, por qué es razonable creer que el yo es una ilusión, y segundo, cómo se genera tal ilusión. Hay varios argumentos posibles a favor de la idea de que el yo es una ilusión, pero el importante, en nuestro caso, es que parece la apuesta naturalista más recomendable, una vez descartado que podamos considerar al animal humano, o el cuerpo, como “a thinking intelligent being, that has reason and reflection, and can consider itself as itself, the same thinking thing, in different times and places”. Si queremos ser naturalistas, y hay buenas razones para serlo, parece una buena idea sostener que no hay tal “ser inteligente que puede considerarse a sí mismo como él mismo”. Lo único que hay son cerebros, o cuerpos que cuentan con cerebros, y lo demás es una ilusión.

¿Cómo se genera esa ilusión? Esta es la parte más difícil en el desarrollo del ficcionalismo. De momento, hay algunas especulaciones sobre cómo puede

sucedan. McDermott (2001), por ejemplo, sostiene que cualquier sistema representacional complejo que pueda representarse a sí mismo, esto es, que tenga una o varias representaciones que son acerca del propio sistema, se representará a sí mismo como una unidad. En esta misma línea va Metzinger (2003, 2009, 2010). A decir de Metzinger, el yo es la manera, o el modo, en que el sistema dotado de auto-representación, se presenta ante sí mismo.

Según Metzinger, el cerebro reconstruye el mundo en la consciencia para hacerlo accesible. Dota a la realidad de coherencia y unidad, pero al precio de generar un modelo que necesariamente confundimos con la propia realidad, ya que el cerebro no puede descubrir que se trata de un modelo –y no de la propia realidad-. El color rosado de la incipiente puesta del sol, nos dice, no es una propiedad del cielo del atardecer, sino del modelo generado por el cerebro. Sin embargo, insiste, el sistema que genera ese modelo no puede tomarlo como tal. Metzinger llama a este tipo de representaciones del mundo –representaciones que el sistema no puede descubrir que lo son– “transparentes”. Digamos que el cerebro nos engaña y nos hace ser realistas *naïf*.

El cerebro humano no se limita a representar el mundo por medio de un modelo fenoménico que confunde con la realidad. También genera un *self-model*. Este segundo es un modelo del propio sistema, no sólo del cerebro, sino del organismo en general. El *self-model* tiene las mismas características que tiene el modelo del mundo: es útil para el propio sistema porque recrea unidad y coherencia donde no las hay, y es transparente en el sentido antes mencionado: no hay forma en que el sistema sepa que el *self-model* no es más que una representación. Contar con un *self-model*, además, hace las interacciones con el mundo y las internas del sistema más fluidas, y dota al sistema de mayor capacidad de control: “It is a particularly user-friendly interface, allowing a biological organism to direct its attention to a critical subset of its own global properties” (2009, p. 106). Según Metzinger, la perspectiva de primera persona, la capacidad de pensar en uno mismo como sí mismo, surge de la integración del *self-model* dentro del modelo del mundo. Pero el yo que se así proyecta no es más que el contenido del *self-model*, i.e., el contenido de una representación fenoménica que no es sino el modo más adecuado que el cerebro ha encontrado de facilitarse las cosas.

Es cierto que no podemos prescindir de creernos un yo y autoatribuirnos una perspectiva fuerte de primera persona, sea en nuestras expresiones lingüísticas o en nuestros pensamientos (Baker, 2013). Sin

embargo, de la inviabilidad de la eliminación en el plano descriptivo no se sigue la eliminación de la entidad presuntamente denotada en el plano ontológico. Esto es, el ficcionalismo, en principio, es una postura viable e inteligible, aunque genere perplejidad. Por otra parte, se trata de una postura relativamente novedosa en el debate mente/cuerpo tradicional. Es, o aspira a ser, un tipo de eliminativismo (ver más abajo), pero no es el tipo de eliminativismo clásico, el del estilo de Skinner (1990, p. 1209): “There is no place in a scientific analysis of behavior for a mind or self.” No está claro que posturas como la de Metzinger sostengan que el yo, por mucho que sea una ilusión, no juegue un papel en la explicación del comportamiento. De hecho, el que el sistema se conciba, o se represente, a sí mismo como lo hace tiene un papel muy importante en la interacción del propio sistema con el mundo y con otros sistemas. Ya hemos hablado de la funcionalidad de la creencia en el yo. Además de lo dicho, cabe mencionar la diferencia de comportamiento que se observa entre individuos que mantienen esa supuesta ilusión y los que, como los pacientes con síndrome de Cotard, que creen que no existen.

Tampoco se trata de una postura reduccionista. El reduccionismo sostiene que el yo no es más que el cerebro, la activación de ciertos patrones de neuronas, el animal entero, o lo que fuere. El “animalismo” de Olson es un reduccionismo. También lo son, aparentemente al menos, las posturas de muchos neurocientíficos, como Crick (1994, p. 3): “You, your joys and your sorrows, your memories and your ambitions, your sense of personal identity and free will, are in fact no more than the behavior of a vast assembly of nerve cells and their associated molecules”. Sin embargo, el ficcionalismo no sostiene que el yo sea idéntico a ninguna otra cosa. La sensación de ser uno mismo tiene un sustrato material, pero no es, como dice Crick, simplemente ese sustrato material. Finalmente, tampoco parece que el ficcionalismo sea un emergentismo, pues el emergentismo típicamente sostiene que lo emergente tiene el mismo rango de existencia que aquello de lo que emerge, aunque la línea entre emergentismo y ficcionalismo sea fácil de cruzar a la hora de ponerlas en palabras. Así, Ramachandran (2007), que es un ficcionalista, describe su postura de este modo: “it now seems increasingly likely that the self is not a holistic property of the entire brain; it arises from the activity of specific sets of interlinked brain circuits”. Sin embargo, la diferencia es grande. El ficcionalismo sostiene que lo que emerge es una ilusión; el emergentismo nos dice que lo emergente es un yo.

El problema del ficcionalismo deriva de su compromiso con los *qualia*. Como hemos expuesto, Metzinger explica la emergencia de la ilusión del yo mediante el contenido cualitativo de una auto-representación (el *self-model*) del sistema de procesamiento de información. El hecho de que tal auto-representación sea transparente, i.e., imperceptible para sí misma, es lo que impide al modelo representarse *qua* modelo propio del sistema. A cambio, se representa fenoménicamente *qua* estable integrador personal de experiencias. De este modo, la propiedad de concebirse a sí mismo como portador de perspectiva egocéntrica y de atribuirse referencia de primera persona, no remite a un nuevo tipo de entidad psicológica irreducible. Todo lo que existe, en su lugar, es una representación del sistema que se presenta a sí misma, no como lo que realmente es (un *self-model*), sino como sujeto personal a través del contenido cualitativo de “sentirse-como-ego-unitario-de-estados-mentales”.

El problema entonces es el siguiente. Para explicar la formación de la “ilusión” del yo, Metzinger recurre a un realismo sobre *qualia*. Las experiencias, incluida la experiencia del yo, no se agotan en su contenido representacional, ni son conceptos fenoménicos, etc., sino que son propiedades irreducibles, modos genuinos en que el cerebro nos presenta la realidad. Según Metzinger, son recreaciones que el cerebro hace de la realidad (el cerebro no nos engaña sólo acerca de sí mismo, sino también acerca del mundo). Sin embargo, recordemos que la admisión de estados mentales con contenido cualitativo requiere, por (PSE), la correspondiente admisión de un sujeto. Los *qualia* no son simplemente modos de representación de la realidad, sino que son modos en los que la realidad se presenta a un sujeto (de hecho, cabe argumentar que el contenido de la experiencia contiene un “for-me-ness” o carácter subjetivo: Kriegel, 2006; Jorba y Vicente, 2014). El sujeto designado por filósofos naturalistas como Metzinger puede ser o bien el yo ficticio o bien el cerebro (o el organismo entero). Supongamos que es el cerebro: el problema entonces, como hemos visto, es cómo hacer compatible esa idea con el principio de la unidad de la experiencia (PUE). El cerebro como un todo puede cesar de existir sin que lo hagan todos y cada uno de los estados mentales conscientes (mientras que *ninguno* de estos estados mentales puede existir sin el sujeto). Esto es tanto como decir que tales estados mentales no dependen unitariamente del cerebro como lo hacen de mí y, por (PUE), que el cerebro no puede figurar como legítimo sujeto de los mismos.

La segunda vía, es decir, sostener que el sujeto de la experiencia es el supuesto yo ficticio, implica reconocer la existencia de ese yo. Si las



experiencias, o los *qualia*, no son reducibles a ninguna otra entidad y tampoco son eliminables, entonces el sujeto de esas experiencias tampoco parece ser reducible (y de hecho, no lo es) ni, más interesante, eliminable. Dado que la postulación de un sujeto es demandada por la postulación de *qualia*, el ficcionalista se ve compelido, o bien a renunciar a los *qualia* y, con ellos, a su explicación naturalista de la emergencia del yo, o bien a resignarse y reconocer la existencia de un portador de los estados mentales causados por el cerebro que, empero, es distinto del cerebro.

La concesión crucial, que ningún naturalista estricto debería estar dispuesto a hacer, es la aceptación de un dominio de estados genuinamente mentales. Si un materialismo es eliminativista, ha de serlo *tout court*. Pero, como remarca Chalmers (1996) el eliminativismo sobre las cualidades subjetivas de las experiencias conscientes no es una posición razonable, aunque sólo sea porque estamos directamente familiarizados con ellas.

Quien, dadas las razones ya expuestas, no esté dispuesto a ceder a las variantes reduccionistas y eliminativistas del materialismo sobre personas y que, simultáneamente, desee evitar el dualismo interaccionista cartesiano, puede aún valorar una versión de la teoría de la constitución material que respete la naturaleza atómica de las personas requerida por (PUE). Una vía transitable podría radicar en aducir que la propiedad adscrita a las personas “ser sujeto unitario de experiencias” y la propiedad adscrita a los cuerpos que las constituyen “estar compuesto por partes” están excluidas de la herencia bidireccional propia de la relación de constitución. Esta táctica tiene la virtud de preservar una concepción asumible de los sujetos personales como portadores unitarios de experiencia al tiempo que permite responder a la objeción de Olson de la duplicación de propiedades sin, por ello, generar tensiones recíprocas entre el dualismo de Lowe y el de Baker. Sugerimos que es ésta una propuesta pendiente de ser sopesada con atención.

## 5. CONCLUSIÓN

Tenemos una noción de persona que remite a una entidad cuya esencia es psicológica. El naturalismo tendría que ser capaz de explicar cómo esta entidad no es otra que la entidad material *cuerpo/animal humano*, o bien de sostener coherentemente que la noción de persona no remite a ninguna entidad mundana. La vía del reduccionismo topa con el problema de que las condiciones de persistencia del cuerpo/animal y de la persona son diferentes.

A ese problema, ya clásico, se suma el argumento de la unidad de la experiencia de Lowe, que muestra que el sujeto de la experiencia no puede ser un organismo por no poder ser sujeto unitario de tal experiencia. Este argumento, en principio, también hace daño a una postura no naturalista, pero moderada, como la teoría de la constitución de Baker, a menos que sea capaz de sostener, de una forma razonada, que el préstamo de propiedades no afecta a la estructura mereológica de constituido y constituyente.

Por su parte, la vía presuntamente eliminativista que hemos analizado, el ficcionalismo, no parece realmente ser capaz de mantener que los sujetos, o las personas, no existan. Parece más bien que el ficcionalismo nos explica cómo se generan los *yoes*. Los sujetos, según el ficcionalismo, no estarían en el mundo natural, o mejor, en el mundo que nos describe la “imagen científica”. Sin embargo, el glosario de entidades de la realidad tendría que incluir tanto a los *qualia* como a los sujetos instanciadores de éstos. En definitiva, parece que el ficcionalismo termina volviéndose dualismo. Otros eliminativismos, más radicales, carecen de la capacidad explicativa del ficcionalismo, aparte de ser más contraintuitivos. Por tanto, de momento cabe mantener que las personas constituyen un problema para la visión naturalista, que no sabe dónde ubicarlas.

#### AGRADECIMIENTOS

Los autores de este artículo agradecen la inestimable ayuda de Fernando Migura en la confección de este trabajo y la inspiración de E.J. Lowe, quien murió mientras trabajábamos sus escritos. Este es nuestro pequeño homenaje a un gran filósofo. Finalmente, gracias a David Pérez Chico por invitarnos a participar en este primer número de *Análisis*. La financiación para la elaboración de este artículo se debe a los proyectos FFI2010-15717 y FFI2011-30074-C02-02 del MINECO.

*Agustín Vicente*  
*Ikerbasque & Universidad del País Vasco, UPV/EHU*  
*agustin.vicente@ehu.es*

*Adrián Sampedro León*  
*Universidad del País Vasco, UPV/EHU*

BIBLIOGRAFÍA

- BAKER, L. R. (1998), "The first- person perspective: a test for naturalism", *American Philosophical Quarterly*, 35, pp. 327-348.
- BAKER, L. R. (1999), "Unity without identity: a new look at material constitution", *Midwest Studies in Philosophy*, pp. 144-165.
- BAKER, L. R. (2000), *Persons and bodies: a constitution view*, Cambridge, Cambridge University Press.
- BAKER, L. R. (2002), "On making things up: constitution and its critics", *Philosophical Topics*, 30, pp. 31-51.
- BAKER, L. R. (2013), *Naturalism and the first-person perspective*, Oxford, Oxford University Press.
- CHALMERS, DAVID J. (1996), *The conscious mind*, Oxford, Oxford University Press.
- CRICK, F. H. C. (1994), *The Astonishing hypothesis: The scientific search for the soul*. New York: Scribner's Sons.
- JORBA, M. Y VICENTE, A. (en prensa), "Cognitive Phenomenology, Access to Contents and Inner Speech", *Journal of Consciousness Studies*.
- KIM, J. (1998), *Mind in a Physical World*, Cambridge, MA: MIT Press: A Bradford Book.
- KRIEGEL, U. (2006), "Consciousness: Phenomenal Consciousness, Access Consciousness, and Scientific Practice", en P. Thagard (ed.), *Handbook of Philosophy of Psychology and Cognitive Science*, Amsterdam: North-Holland, pp. 195-217
- LOCKE, J. (1690), *An essay concerning human understanding*, version abreviada (1984), Glasgow, Collins Fount Paperbacks.
- LOWE, E.J. (1996), *Subjects of experience*, Cambridge, Cambridge University Press.
- LOWE, E.J. (2009), *Dualism*, en MCLAUGHLIN, BRIAN P.& BECKERMANN, ANSGAR& WALTER, SVEN (eds.), *The Oxford Handbook of Philosophy of Mind*, Oxford, Oxford University Press, pp. 66-84.

- LYCAN, WILLIAM G. (2013): “Is property dualism better off than substance dualism?”, *Philosophical Studies*, 164 (2): 533-542.
- MCDERMOTT, D. (2001): *Mind and Mechanism* MIT Press.
- METZINGER, T. (2003): *Being No One. The Self-Model Theory of Subjectivity*. Cambridge, MA: MIT Press.
- METZINGER, T. (2009): *The ego tunnel. The science of the mind and the myth of the self*, New York, Basic Books.
- METZINGER, T. (2010): “The no-self alternative”, en GALLAGHER, SHAUN (ed.) *The Oxford Handbook of the Self*, Oxford, Oxford University Press, pp. 279-296.
- OLSON, ERIC T. (1998): “Human atoms”, *Australasian Journal of Philosophy*, 76, pp. 396-406.
- OLSON, ERIC T. (2007): “What are we?”, *Journal of Consciousness Studies* 14, pp. 37-55.
- RAMACHANDRAN, V. S. (2007). “The neurology of self-awareness”, [1.8.07] *Edge* ([www.edge.org](http://www.edge.org)).
- ROBINS, R. W., TRACY, J. L., & TRZESNIEWSKI, K. (2008), “Naturalizing the self”, en O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of Personality* (3 rd Edition), pp. 421-447.
- SKINNER, B. F. (1990), “Can psychology be a science of the mind?” *American Psychologist*, 45, pp. 1206-1210.
- STRAWSON, P.F (1959), *Individuals*, London: Methuen.
- WIGGINS, DAVID (1968), “On being in the same place at the same time”, *Philosophical Review*, 77, pp. 90-95.
- WIGGINS, DAVID (1980), *Sameness and substance*, Cambridge, Cambridge University Press.
- ZIMMERMAN, DEAN W. (2010), “From property dualism to substance dualism”, *Proceedings of the Aristotelian Society*, Supplementary Vol. 84, pp. 119-150.