



Artículo de revisión

La Pieza China: ¿un experimento mental con sesgo cartesiano?

The Chinese Room: A thought experiment with a Cartesian flavor?

Rodrigo González^{1*}

1 Departamento de Filosofía y Centro de Estudios Cognitivos, Facultad de Filosofía y Humanidades, Universidad de Chile. Santiago, Chile.

Resumen

Este ensayo examina un experimento mental clásico de John Searle en filosofía de la mente, cuyo argumento ha sido descalificado por Dennett y Hofstadter como una bomba de intuiciones no confiable. Lo que se defiende aquí es que este experimento mental tiene un sesgo cartesiano, pero ello no obsta para que no sea confiable. En efecto, la característica principal de la Pieza China es depender de un agente cognitivo consciente que realiza el experimento y, en particular, de quien no se puede prescindir para saber si ocurre o no entendimiento lingüístico luego del procesamiento de símbolos. En efecto, el punto de vista de la 1ª persona de quien lleva a cabo el experimento es, al igual que en otras instancias, lo que permite el testeo de una hipótesis sobre la naturaleza de lo mental.

Palabras clave: la Pieza China, experimento mental, bomba de intuición, cartesianismo dualista

Abstract

This article examines a classical thought experiment by John Searle in philosophy of mind, which Dennett and Hofstadter have accused as a non reliable intuition pump. As I argue here, although this thought experiment involves a Cartesian bias, this does not mean it is unreliable. Indeed, the Chinese Room most important feature is that it depends upon a conscious cognitive agent who performs the experiment and, in particular, someone who cannot be jettisoned in order to know whether or not linguistic understanding takes place after symbol processing. Certainly, the first person viewpoint of the agent who carries out the experiment is, like in other instances, what allows one to test a hypothesis about the nature of the mind.

Keywords: Chinese Room, thought experiment, intuition pump, Cartesian dualism

* Correspondencia: rodgonfer@gmail.com. Departamento de Filosofía y Centro de Estudios Cognitivos, Universidad de Chile, Avenida Ignacio Carrera Pinto 1025, Ñuñoa, Santiago de Chile. Teléfono: +56 (2) 9 787 142, +56 (2) 9 787 004
Recibido: 13-09-11. Revisión desde: 22-11-11. Aceptado: 28-12-11

Introducción

Buena parte de la filosofía de la mente en la década de 1980 y también posterior se dedicó a examinar la Pieza China de John Searle. Su creador la postuló en 1980 como una manera *a priori* de saber si el procesamiento simbólico informacional era suficiente para el entendimiento lingüístico, cuestión que contaría como evidencia empírica de que la mente es un computador programado.

Sin temor a equivocarse, uno puede afirmar que el experimento mental de la Pieza China polarizó a la comunidad científica y filosófica. Nadie quedó indiferente frente al experimento mental, al punto de que suscitó odios y simpatías por igual. Algunos la han calificado de falaz e irrelevante; otros piensan que es una situación hipotética que describe de manera fiel qué implica computar símbolos sobre la base de reglas formales.

Searle mismo aprovechó el experimento mental no solo para manifestar su juicio sobre los computadores digitales, sino además para comenzar a forjar su propia teoría sobre la mente, la que posteriormente denominó “naturalismo biológico” (Searle, 1992). Puesto de manera simple y económica, esta teoría sostiene que la mente es como la digestión o la fotosíntesis: un proceso biológico sujeto a leyes y, como tal, de naturaleza causal.

Una consecuencia de lo anterior directamente relacionada con la inteligencia artificial (IA) es que, según Searle, replicar la mente requiere duplicar los mismos poderes causales del cerebro. Esta postura se opone a Descartes, quien sostiene que lo mental es de una naturaleza diferente de lo material y espacial, al punto que ningún mecanismo es capaz, *en principio*, de tener la flexibilidad de la razón. De hecho, el siguiente pasaje de Descartes inspiró a la comunidad de la IA a rebatirlo:

“Pero (una máquina) no podría ordenar las palabras de formas diferentes para responder al significado que se dice en su presencia, como incluso el menos inteligente de los humanos puede hacer (...) Incluso, aunque tales máquinas podrían hacer algunas cosas tal como nosotros las hacemos, o de mejor forma, inevitablemente fallarían en otras, lo cual revelaría que no están actuando de acuerdo con su entendimiento, sino por la pura disposición de sus órganos” (Descartes, 1994 [1637], p. 113).

De este modo, Descartes argumenta que las máquinas están imposibilitadas en principio de ser racionales por carecer de lenguaje y ser mecánicamente inflexibles en su conducta. Searle, por el contrario, cree que el hombre es una máquina con poderes causales suficientes para tener estados mentales, los cuales serían otorgados por el cerebro y el sistema nervioso, no por programas computacionales al estilo de la inteligencia artificial clásica, esto es, aquella que postula que el cerebro es la implementación de una máquina programada.

Pese a la notable diferencia entre las filosofías de Descartes y Searle, el experimento mental de este último posee un sesgo cartesiano producto de la preeminencia que adquiere el agente cognitivo o el ejecutante-experimentador. En este trabajo examino tal sesgo, mostrando que no es privativo de este experimento mental, sino que se da en todos aquellos que exploran el contraste entre la propia mente y lo descrito por alguna teoría filosófica.

Naturalmente, tal contraste se funda en la introspección y la conciencia, dos instancias que no parecen confiables a todos

los filósofos de la mente. Hofstadter y Dennett (1981) acuñan el término “bomba de intuición” para aquellos experimentos mentales que, describiendo una historia pintoresca, giran las perillas de la mente del propio experimentador para forzarlo a extraer una conclusión. En este sentido, la Pieza China sería un ejemplo paradigmático de bomba de intuición. En vista de su crítico análisis, propongo que el experimento mental de Searle, al igual que otros, es confiable porque se basa en el punto de vista de la 1ª persona, lo cual contrasta notoriamente con el procesamiento algorítmico de una máquina programada que, por definición, no tiene ningún punto de vista. Es decir, hay una manera de ser como un organismo, mientras que no hay manera de ser como una máquina y este contraste representa un desafío no solo para la inteligencia artificial, sino para todas las teorías materialistas.

El trabajo se divide en tres secciones. En la primera reviso el experimento mental de la Pieza China, proporcionando el contexto para entender su alcance y sesgo. En la segunda analizo la crítica de Hofstadter y Dennett (1981), quienes catalogan este experimento mental como una bomba de intuiciones. En la sección final explico por qué Searle desdeña el dualismo cartesiano y muestro que, pese a la tradicional interpretación a la que da lugar, no implica descartar de plano factores materiales involucrados en causar la mente.

El experimento mental de la Pieza China

Teniendo en cuenta el juicio de Descartes sobre el lenguaje y los mecanismos, la IA clásica del siglo XX se enfocó en mostrar cómo es posible que una máquina programada sostenga conversaciones como un humano. Turing (1950) aceptó el desafío del filósofo francés, por lo que intentó reemplazar la pregunta: “¿pueden pensar las máquinas?” por el juego de la imitación. Este contempla un conjunto de jueces que deben ser engañados por una máquina programada mediante la imitación que esta hace de la capacidad lingüística de un hablante en el lapso de una conversación típica de cinco minutos. Si un computador digital es capaz de engañar al menos a 30% de los jueces, esto, según Turing, indicaría evidencia inductiva suficiente para suponer que la máquina programada tiene estados mentales tal como un cerebro los tiene¹.

Varios programas se diseñaron para pasar el Test de Turing, especialmente en la década de 1970. En particular, Schank y Abelson (1977) propusieron su programa SAM, *Script Applier Mechanism*, como una manera de reducir el entendimiento de una lengua a la descomposición recursiva de un programa computacional. Como es sabido, tal descomposición opera a través de la simplificación de operaciones complejas, esto es, “entender”, subsumiéndose en operaciones cada vez más simples y mecánicas: hacer calzar historias con *scripts*. Así, SAM funciona gracias a una base de datos que contiene secuencias de historias o *scripts*, y el programa las combina de modo tal que extrae información que no esté necesariamente explícita. Tal capacidad combinatoria de historias permite al programa responder preguntas relevantes acerca de una historia en particular.

Por ejemplo, si se pregunta si alguien dejó propina luego de haber recibido un plato de pizza carbonizada en un restaurante, el programa responderá que el cliente no dejó nada, y tal

¹ Para un examen crítico del Test de Turing y las usuales interpretaciones a las que da lugar, sugiero González (2007).

vez agregará que este pidió el libro de sugerencias y reclamos para narrar su mala experiencia. Ciertamente, ambos hechos no se infieren de la desafortunada experiencia del cliente, pero sí impactan como altamente probables a quien frecuente restaurantes y entienda la cadena de eventos que usualmente ocurre en el *script* "restaurante". De tal modo, Schank y Abelson reducen el entendimiento lingüístico a la capacidad de hacer calzar y combinar diferentes historias de manera mecánica o algorítmica, todo lo cual es suficiente para pasar el Test de Turing.

En vista de la tesis de Schank y Abelson, Searle (1980, p. 67) propone la distinción entre inteligencia artificial fuerte y débil. De acuerdo con la IA débil, SAM simplemente simula entendimiento, en el sentido de que la manipulación simbólica de historias no permite afirmar que el programa computacional es, como cuestión de hecho, una mente con estados intencionales. Así, la IA débil es la actitud del investigador cauto e instrumentalista: las teorías y modelos de la IA permiten una mejor comprensión de qué es la mente, pero no puede sostenerse que lo descrito por la teoría es *literalmente* una mente. Puesto de manera breve y simple: para la IA débil, SAM simularía y tendría conducta lingüística como si fuera una mente, lo que permite una mejor comprensión de cómo podría funcionar ésta.

En cambio, según la IA fuerte es posible afirmar que SAM entiende las historias tal como cualquier persona lo hace. Si esto es así, entonces se dice que SAM poseería estados mentales, en tanto tendría creencias respecto de las historias y las preguntas que se siguen de estas. En palabras simples de Schank y Abelson, su programa sería una mente, pues operaría de la misma manera como la mente de cualquier persona lo hace, esto es, mediante la manipulación formal de representaciones simbólicas que combinan historias cuando uno se enfrenta a una serie de preguntas.

Dada la IA fuerte, Searle postula un experimento mental que, en sus palabras, sería definitivo para probar su falsedad: la Pieza China. Esta es la descripción de una situación hipotética en que él, un hablante nativo de inglés que no sabe chino, está encerrado en una pieza y debe contestar preguntas acerca de una historia en tal idioma. A través de la hendidura de una puerta recibe ideogramas chinos, los cuales claramente no entiende por no conocer una palabra del idioma. Recibe tres conjuntos de ideogramas. El primero es de los *scripts* o secuencias de eventos típicos. El segundo representa la narración de una historia particular en chino. Y el tercero es un conjunto de preguntas sobre la historia narrada.

Lo sorprendente es que Searle, pese a no entender una palabra de chino, es capaz de manipular símbolos en función de un libro de reglas, el cual le dicta qué hacer con los ideogramas solamente con base en su forma. Nuevamente, Searle no entiende los símbolos y, como tales, son solo complejas líneas y trazados, sin significar nada para él. Aunque no hay entendimiento por parte de quien ejecuta el experimento mental, Searle en este caso, sería posible que respondiese a las preguntas sobre la historia de manera coherente y relevante. En efecto, si factores técnicos como el tiempo y la velocidad no tuvieran importancia, sería posible en principio que Searle tomara el libro de reglas y manipulara los tres conjuntos de ideogramas chinos teniendo en cuenta solo su forma, con lo cual podría elaborar respuestas también en este idioma.

Si un grupo de observadores externos a la pieza, los mismísimos chinos que entregan los ideogramas a Searle, recibieran respuestas de este gracias a la manipulación simbólica

que permite el libro de reglas, dichos observadores quedarían sorprendidos por las respuestas. Más aún, les daría la impresión de que existe el entendimiento lingüístico de un hablante nativo, con lo cual se pasaría el Test de Turing.

Una cuestión que ha sido enfatizada por Searle es la simplicidad de qué significa "entender", una actividad que se contrasta con la manipulación simbólica. De hecho, afirma lo siguiente:

"En el caso del inglés entiendo todo, y no existe razón alguna para suponer que mi entendimiento tenga que ver con el programa de computador, esto es, con las operaciones computacionales basadas en elementos especificados de manera puramente formal. En tanto que el programa se define en términos de operaciones computacionales sobre elementos definidos de modo puramente formal, lo que el ejemplo [de la Pieza China] sugiere es que estos mismos no tienen una conexión interesante con entender [...]"

"Hay ejemplos claros de 'entender' que literalmente se aplican y casos claros en que no se aplican; y estos dos tipos de casos son todo lo que se necesita para el argumento. Entiendo historias en inglés; en grado menor entiendo historias en francés; en grado aún menor, historias en alemán; y en chino, en modo alguno. Por otra parte, mi automóvil y calculadora no entienden nada: no están pensadas para este tipo de actividad. A menudo se le atribuye "entendimiento" y otros predicados cognitivos mediante metáfora a automóviles, calculadoras y otros artefactos, pero nada se prueba mediante esas atribuciones" (Searle 1980, pp. 70-72).

Sistematizando más formalmente su argumento, en otro artículo remarca la misma idea de la siguiente manera:

"El punto del experimento mental es este: si no entiendo chino solamente sobre la base de correr un programa para entender chino, entonces no lo hará ningún computador digital sobre esas bases. Los computadores digitales solamente manipulan símbolos formales en función de las reglas de un programa" (Searle 1990, p. 20).

El significado del experimento mental es, pese a sus críticos, claro. Si alguien corre un programa para entender chino, la persona en cuestión será incapaz de entender qué significan los ideogramas solo en función de la manipulación formal, sintáctica, de los símbolos de carácter representacional. En consecuencia, qué representan los símbolos está vedado para quien opera en función de formas y reglas.

Ahora bien, lo crucial respecto de este tipo de experimentación mental es que "entender" es una actividad evaluada por quien ejecuta el experimento mental mismo. En el caso de la Pieza China, se evalúa la verdad o falsedad de la IA fuerte gracias a quien aplica los principios de esta a su propia mente, y así determina qué sucedería si esta operase según lo que dicta aquella. En efecto, la mera manipulación de símbolos a causa de su forma, no de su contenido, es insuficiente para entender historias en chino, porque dicha manipulación opera mediante pura sintaxis y como tal no implica semántica, i.e., "saber" respecto del contenido de los símbolos.

Las réplicas y objeciones a la Pieza China de John Searle son numerosas. De hecho, el mismo considera un buen número

de estas en 1980 (pp. 72-87)². Sin embargo, en lo que sigue me concentro en una objeción particular de Hofstadter y Dennett (1981), quienes afirman que este tipo de experimentación mental es tan fascinante como engañosa, al contar como una mera bomba de intuiciones.

¿Una bomba de intuiciones no confiable? El rol crucial del agente cognitivo que ejecuta el experimento mental

Es importante aclarar a qué se refieren Hofstadter y Dennett (1981) con “bomba de intuición”. Según ellos, este apelativo puede emplearse contra una clase de experimento mental que mueve a la persuasión al narrar una historia pintoresca que, por lo general, involucra girar las perillas de la propia mente del experimentador. Así, se alude a un escenario imaginario en que el propio yo de la mente es sometido a la simulación de una actividad mental, y esto presiona al experimentador a obtener una conclusión. En particular, afirman lo siguiente:

“Cada miembro de esta clase de experimento mental se define por una elección particular de “vueltas de perillas” en el generador del experimento mental mismo. Su propósito es crear —en el yo de su mente— varias simulaciones imaginarias de la actividad mental humana. Cada experimento mental es una bomba de intuición (término de Dennett) que magnifica una faceta u otra del problema, tendiente a empujar al lector a ciertas conclusiones” (Hofstadter y Dennett, 1981, p. 375).

A las perillas de toda bomba de intuición, Hofstadter y Dennett agregan otro elemento: el “demonio del experimento mental”. Este es quien ejecuta el procedimiento experimental que supuestamente puede apresurar la conclusión al tener un punto de vista privilegiado respecto de la operación descrita. En el caso de la Pieza China, el demonio sería quien ejecuta el experimento mental y, por tanto, realiza operaciones computacionales imaginarias sobre la base del libro de reglas. Esto, según Hofstadter y Dennett, haría a la Pieza China un experimento mental no confiable. De hecho, según ellos, es perfectamente posible que ocurra entendimiento en la pieza misma gracias a la manipulación simbólica. Esta posibilidad, muy popular entre los defensores de la IA fuerte, se denomina “la réplica del sistema”. Básicamente, establece que aunque Searle no entiende chino existe entendimiento en la pieza en su totalidad, el cual es inalcanzable para el demonio ejecutante del experimento. Insisten dichos defensores que la Pieza China, el sistema total, pasaría el Test de Turing.

Searle justamente elabora una respuesta a la réplica del sistema basada en otro escenario imaginario: una situación tal en que el ejecutante es capaz de introducir mentalmente todos los elementos de la pieza, incluyendo el libro de reglas que memoriza, y es de ese modo capaz de responder las preguntas, aunque nuevamente sin comprender chino. Nada hay en la pieza que no esté en el experimentador-ejecutante, y aun así no habría entendimiento lingüístico. Pese a que esta nueva réplica de Searle ha sido criticada como en extremo extravagante y técnicamente imposible —una razón más para considerar la Pieza China como una bomba de intuición— tiene plausibilidad desde el punto de vista lógico. Efectivamente, que la Pieza

China sea técnicamente imposible no implica que sea lógicamente imposible, y así que no se siga la conclusión que Searle extrae, a saber, que la pura manipulación sintáctica no necesariamente implica semántica y el consiguiente entendimiento lingüístico.

De hecho, una manera de apoyar la argumentación de Searle es imaginar una situación en que se determinen qué factores del experimento mental son relevantes o no. Por ejemplo, es innecesario que haya una rendija en la puerta, que las reglas se establezcan en un libro, que haya mobiliario en la pieza, y así sucesivamente. Existe una serie de elementos en la pieza que son irrelevantes para corroborar de qué manera el entendimiento no emerge de la sintaxis. Incluso, pudiera ser que la pieza misma no existiera y que, como Searle sugiere, se corriera el programa en la propia mente. Sin embargo, hay un elemento que es fundamental y sin el cual la operación del experimento mental resulta imposible: la persona que contrasta la manipulación simbólica con lo que usualmente ocurre cuando se entiende una lengua. Sin este elemento no hay experimento mental, y ese punto es lo que pasan por alto Hofstadter y Dennett. En efecto, debe haber algún experimentador-ejecutante que evalúe si efectivamente hay entendimiento lingüístico y esto, desde mi perspectiva, haría que este experimento mental pudiese considerarse como cartesiano. ¿En qué sentido?

El punto de vista de la 1ª persona del agente cognitivo es lo que permite contrastar la manipulación simbólica con el entendimiento genuino, cuestión crucial al ejecutar el experimento. Solo el agente o ejecutante-experimentador *sabe y está consciente de* cómo es el entendimiento de una lengua, el cual se diferencia notoriamente de la operación algorítmica de manipular símbolos. Así, se contrastan dos puntos de vista: el de la 1ª persona del experimentador, que sabe que es ser como una mente cuando esta entiende, y el de una máquina programada, que por definición no tiene ningún punto de vista. El punto de vista de la 1ª persona es el sesgo cartesiano de la Pieza China que Searle, por razones que se analizarán en la próxima sección, evitaría a toda costa calificar de tal modo.

Sin embargo, la singularidad de la Pieza China no es privativa de este experimento mental solamente. Existen al menos dos experimentos mentales que, al igual que la Pieza China, también han sido descalificados como bombas de intuiciones: el murciélago de Nagel (1974) y *Mary*, la neurocientista de Frank Jackson (1986).

En el primero se invita al lector a considerar qué sucedería si se tuviera conocimiento del aparato neurofisiológico de un murciélago. Aunque un agente cognitivo supiera a ciencia cierta *cómo* conoce un murciélago, es decir, tuviera una teoría que describiera de qué manera el murciélago procesa información sensorial, ello no sería suficiente para imaginar qué es tener la experiencia consciente de ser tal mamífero. Dado que todo el material de la imaginación es proporcionado por la experiencia y esta es producto del aparato neurofisiológico, para quien ejecuta el experimento mental no es posible *saber ni imaginar* que es ser como un murciélago.

Por otra parte, el experimento mental de *Mary* la neurocientista apunta a algo similar. Jackson invita al lector a considerar qué sucedería si *Mary* estuviese toda su vida encerrada en una habitación en blanco y negro. Pese a lo anterior, *Mary podría saber* todo lo que hay que saber acerca del color a través de lo que lee en sus libros de texto. Con todo, si un día saliera de la pieza en blanco y negro, sin duda aprendería algo nuevo

² Para una revisión de las defensas y objeciones más importantes, véase Preston y Bishop (2002).

acerca de los colores, en tanto llegaría a estar consciente de qué son estos.

Tanto en la Pieza China como en estos experimentos mentales hay preponderancia del punto de vista del agente cognitivo que ejecuta el experimento, y son en el mismo sentido cartesianos, pues en ellos se evalúa la conclusión sobre la base de los propios estados conscientes de un experimentador. En todos participa la conciencia fenoménica, la cual permite contrastar “qué es ser como” dicho agente y lo descrito por teorías específicas de qué es la mente.

En el caso de *Mary*, el agente cognitivo evalúa qué es ser como alguien que sabe sobre los colores, pero nunca ha tenido la experiencia visual de estos. Con el murciélago de Nagel se examina la imposibilidad de imaginar la experiencia consciente de ser otro organismo desde la perspectiva de un humano. El fundamento central es que un humano solo puede tener la experiencia consciente propia. Y, finalmente, en la Pieza China se explora cómo un experimentador-ejecutante es incapaz de estar consciente de la experiencia asociada a entender chino cuando implementa un programa para tal efecto.

Tanto la Pieza China como el murciélago de Nagel y *Mary* permiten el testeo filosófico de hipótesis acerca de la naturaleza de la mente. Esto no involucra bombear puras intuiciones, sino comparar estados cognitivos conscientes con lo que sostienen algunas teorías sobre la mente. Por ejemplo, la Pieza China testea qué sucedería si la mente operase como un computador digital para responder preguntas sobre historias. El murciélago de Nagel permite explorar qué sucedería si el fisicismo fuera verdadero, esto es, permite determinar si una descripción física de la experiencia consciente es suficiente para comprender qué es ser como ese organismo. Y *Mary* también ayuda a evaluar qué sucedería si el fisicismo fuese la teoría correcta de la mente, es decir, permite ponderar qué ocurriría si uno supiese todo lo que es posible saber teóricamente de la experiencia visual, incluso si nunca se hubiera tenido.

Pero, tal como se describió, Searle no está dispuesto a aceptar que su experimento mental tiene un sesgo cartesiano, otorgando certeza de qué es la mente gracias a la preeminencia del agente cognitivo que ejecuta el experimento. Este filósofo ha condenado el cartesianismo en diversas ocasiones, al igual que la mayor parte de la tradición en filosofía de la mente contemporánea.

La tradicional interpretación de Searle sobre el dualismo de Descartes

¿Por qué Searle no está dispuesto a aceptar que su experimento mental explota la preponderancia de un agente cognitivo consciente, y que se funda en la conciencia fenoménica para establecer que la IA fuerte es imposible? La respuesta se vincula con la usual caracterización que hace Searle de la filosofía cartesiana.

Aunque Searle le atribuye a Descartes haber inaugurado la filosofía de la mente, su opinión de este último es totalmente negativa, incluso responsabilizándolo de heredar el problema mente-cuerpo a la tradición. Por ejemplo, en la sección de su libro “La mente: una breve introducción”, intitulada *Descartes y otros desastres*, afirma lo siguiente:

“La doctrina más famosa de Descartes es el dualismo, la idea de que el mundo se divide en dos diferentes clases de

substancias o entidades que pueden existir por sí mismas. Estas son la substancia mental y la substancia física. La versión de Descartes del dualismo es a veces llamada “dualismo de substancia” (...) Al decir que la esencia de la mente es la conciencia, Descartes defiende la idea de que somos la clase de entidades que somos porque estamos conscientes, y que siempre estamos en algún estado consciente, o cesaríamos de existir si cesáramos de estar en un estado consciente (...) Al sostener que la esencia de lo corporal es la extensión, Descartes defiende la idea de que los cuerpos tienen dimensiones espaciales: el escritorio en frente de mí, el planeta Tierra, y el automóvil en el estacionamiento son extensos o figuran en el espacio” (Searle, 2004, p. 9, énfasis en original).

Si bien esta caracterización substancialista es usual respecto del dualismo cartesiano, no deja de ser discutible y ha sido objeto de larga disputa académica. El propio Descartes fue algo ambiguo respecto de la distinción entre lo corporal, la *res extensa*, y lo mental, la *res cogitans*. Y lo fue porque el modo de exploración que emplea en sus “Meditaciones metafísicas” es más bien epistémico, en el sentido de que, al intentar refutar al escéptico, destaca qué es lo que puede saberse con certeza. Por este motivo llega a la idea del *cogito*:

“El pensamiento es un atributo que me pertenece, siendo el único que no puede separarse de mí. *Yo soy, yo existo*; eso es cierto, pero ¿cuánto tiempo? Todo el tiempo que estoy pensando: pues quizás ocurriese que si yo cesara de pensar, cesaría al mismo tiempo de existir. No admito ahora nada que no sea necesariamente verdadero: no soy más que una cosa que piensa, es decir, un espíritu, un entendimiento o una razón. (...) Ahora bien: ya sé de cierto qué soy y que, a la vez, puede ocurrir que todas esas imágenes y, en general, todas las cosas referidas a la naturaleza del cuerpo no sean más que sueños y quimeras” (Descartes, 1977 [1642], pp. 25-26).

Este pasaje parece corroborar lo aseverado por Searle, ya que Descartes indica que la mente es lo más fácil de conocer y que es un atributo que no puede separarse de uno. Incluso, afirma que *quizás* si cesara de pensar, cesaría de existir.

Pero aquí hay que ser particularmente cuidadoso al leer lo argumentado por Descartes: él habla de lo que es necesariamente verdadero, en el sentido de qué es más fácil de conocer como tal. Esto indica nuevamente el modo epistémico de su exploración, de lo que se *sabe*, *no de lo que es*, y por lo mismo no debe apresurarse una conclusión metafísica al respecto. De hecho, dos pasajes más permiten morigerar la interpretación de Searle sobre el dualismo cartesiano.

Primero, en la 6ª Meditación Descartes establece cómo se tiene la idea de unión entre mente y cuerpo:

“Me enseña también la naturaleza, mediante las sensaciones de dolor, hambre, sed, etcétera, que yo no solo estoy en mi cuerpo como un piloto en su navío, sino que estoy tan íntimamente unido y como mezclado con él que es como si formásemos una sola cosa. Pues si ello no fuera así, no sentiría yo dolor cuando mi cuerpo está herido, pues no soy sino una cosa que piensa, y percibiría esta herida con el solo entendimiento, como un piloto percibe, por medio de la vista, que algo se rompe en su nave; y cuando mi cuerpo necesita beber y comer, lo entendería yo sin más, no avisándome de ello sensaciones confusas de hambre y sed. Pues, en efecto, tales sentimientos de hambre, sed, dolor, etcétera, no son sino modos confusos de pensar, nacidos de

esa unión y especie de mezcla del espíritu con el cuerpo, y dependientes de ella” (Descartes, 1977 [1642], p. 68).

La mente se encuentra unida con lo corporal también, y de eso hay experiencia concreta, cuestión que es compatible con *abstraer* lo mental de lo material cuando se explora qué es la mente. Pese a tal capacidad de abstracción, hay evidencia de la interconexión entre mente y cuerpo en el caso de las sensaciones, las cuales se manifiestan con confusos modos de pensar, como el dolor, la sed, etc.

En el segundo pasaje, el propio Descartes no descarta que la mente pueda poseer algo material que no esté reconocido por el entendimiento, que no sea *conocido* como tal. Una carta a Clerselier de 1646 es destacada por Clarke (2003) así:

“Dije en un lugar que mientras que el alma está en duda acerca de las cosas materiales, se conoce *praecise tantum* —“en sentido estricto solamente”— como una substancia inmaterial. Y siete u ocho líneas más abajo mostré que por las palabras “en sentido estricto” no quiero decir una exclusión o negación, sino una abstracción de las cosas materiales. Debido a que dije que a pesar de esto no estamos seguros de que no haya nada corporal en el alma, incluso aunque no pueda reconocerse nada corpóreo en ésta” (Descartes, 1646, en Clarke, 2003).

De esta forma, hay que ser particularmente cuidadoso con la interpretación del dualismo cartesiano. Descartes explora qué puede establecer como cierto sobre la mente y el cuerpo de un modo bastante similar a los experimentos mentales descritos en la sección previa: se compara el estado consciente de un ejecutante-experimentador con otras actividades cognitivas y, en función de esta comparación, se llega a inferir conclusiones. En consecuencia, es posible sostener que los experimentos mentales sobre la propia mente son de índole cartesiana, pues exploran qué hay de cierto sobre la mente al considerar teorías que la explican.

La moraleja de toda la discusión sobre la Pieza China es, entonces, que al momento de considerar qué es la mente se necesita de cierta moderación y espíritu ecléctico. En efecto, se equivocan aquellos que simplifican en extremo lo mental y, por lo mismo, descalifican el experimento de Searle, alegando que es una mera bomba de intuiciones, una falacia incapaz de mostrar algo útil en filosofía de la mente. Por el contrario, la Pieza China, al igual que otros experimentos mentales, indica que no es posible deshacerse del punto de vista de la 1ª persona en cualquier explicación de lo mental. Asimismo, yerran los que simplifican la filosofía cartesiana, sin reparar que el propio Descartes piensa que el cerebro y otros factores materiales podrían tener participación en la explicación de la mente. Los experimentos mentales sobre ésta y entre estos la Pieza China, nos recuerdan que hay mucho trabajo todavía por hacer para explicar la mente, sin descuidar el punto de vista de la 1ª persona que esta involucra. En efecto, la conciencia sigue siendo la piedra de tope de cualquier teoría, y sigue representando un desafío para el filósofo y el científico que se abocan al estudio de lo mental.

Referencias

- Clarke, D. (2003). *Descartes' Theory of Mind*. Oxford: OUP.
- Descartes, R. (1977). *Meditaciones metafísicas*. Madrid: Alfaguara.
- Descartes, R. (1994). *Discurso del método*. Madrid: Alianza.
- González, R. (2007). El Test de Turing: dos mitos, un dogma. *Revista de Filosofía Universidad de Chile*, 63, 37-53.
- Hofstadter, D. & Dennett, D. (1981). *The Mind's I*. New York: Basic Books.
- Jackson, F. (1986). What Mary didn't know. *The Journal of Philosophy*, 83(5), 291-295. Reimpreso en T. O'Connor and D. Robb (Eds.) *Philosophy of Mind: Contemporary Readings*. London: Routledge, 458-463.
- Preston, J. & Bishop, M. (2002). *Views into the Chinese Room*. Oxford: OUP.
- Schank, R. & Abelson, R. (1977). *Scripts, Plans, Goals and Understanding*. Hillsdale, NJ: Erlbaum.
- Searle, J. (1980). Minds, Brains and Programs. *The Behavioral and Brain Sciences*, 3, 417-424. Reimpreso en M. Boden (ed.) *The Philosophy of Artificial Intelligence*. Oxford: OUP, 67-88.
- Searle, J. (1990). Is the Brain's Mind a Computer Program. *Scientific American*, January, 20-25.
- Searle, J. (1992). *The Rediscovery of the Mind*. Cambridge, Mass., MIT Press.
- Nagel, T. (1974). What is it like to be a bat? *Philosophical Review*, 83, 435-450. Reimpreso en D. Chalmers (Ed.) *Philosophy of Mind: Classical and Contemporary Readings*. Oxford: OUP, 219-226.
- Turing, A. (1950). Computing intelligence and machinery. *Mind*, LIX(2236), (Oct. 1950), 433-60. Reimpreso en M. Boden (Ed.) *The Philosophy of Artificial Intelligence*. Oxford: OUP, 40-66.