

DAVID VILLENA SALDAÑA

QUALIA

QUALIA

QUALIA

Resumen

El presente artículo muestra el carácter problemático de los *qualia* para una teoría de lo mental. Con el término ‘qualia,’ hacemos referencia a aspectos no intencionales de estados mentales eminentemente cualitativos como pueden ser las experiencias perceptuales, las emociones, los humores y las sensaciones corporales. Estos aspectos no intencionales suelen caracterizarse como intrínsecos, inefables, infalibles, atómicos, privados, directos e irreducibles a lo físico. Se explica, además, el argumento de los *qualia* ausentes como crítica al funcionalismo sobre lo mental.

Palabras clave: *qualia*; *qualia* ausentes; conciencia; problema mente-cuerpo; funcionalismo.

Abstract

This article shows the problematic nature of qualia for a theory of mind. With the term ‘qualia’ we refer to the unintentional aspects of mental states especially qualitative such as perceptual experiences, emotions, moods and bodily sensations. These unintended aspects are often characterized as intrinsic, ineffable, infallible, atomic, private, direct and physically irreducible. It also explains the argument of absent qualia as a critique of functionalism on the mental.

Keywords: qualia; absent qualia; consciousness; mind- body problem; functionalism.

Résumé

Le présent article montre le caractère problématique des *qualia* pour une théorie sur le mental. Avec le terme 'qualia', nous faisons référence à des aspects non-intentionnels d'états mentaux éminemment qualitatifs comme peuvent l'être les expériences perceptuelles, les émotions, les humeurs et les sensations corporelles. Ces aspects non-intentionnels tendent à être caractérisés comme étant intrinsèques, ineffables, infaillibles, atomiques, privés, directs et irréductibles au physique. De plus, est expliqué l'argument des *qualia* absents comme critique du fonctionnalisme sur le mental.

Mots clés: *qualia*; *qualia* absents; conscience; problème corps-esprit; fonctionnalisme.

Fecha de recepción : 24/07/2016

Fecha de aceptación : 09/09/2016

1. ¿Qué son los *qualia*?

Dadas las condiciones sensoriales estándar para un ser humano cualquiera, la experiencia perceptual de un tomate maduro (rojo) presenta un aspecto cualitativo diferente al de la experiencia perceptual de un trozo de mantequilla (amarillo). Precisarlos podría sonar a trivialidad a oídos de algunos. Habría quien se apresure a preguntar: "¿Es que acaso no se nota?". La certeza que se suele albergar sobre el particular es consecuencia de estar tratando con una opinión introspectiva, opiniones en relación con las cuales consideramos ser infalibles. En efecto, la introspección nos dice que no se siente igual ver rojo que ver amarillo. Si esto es así, ello equivale a afirmar que la experiencia del rojo tiene un *quale* distinto al de la experiencia del amarillo¹. Y lo mismo que ocurre

1 Para la historia del término filosófico 'quale' y su uso primigenio por parte de C. S. Peirce y W. James hasta su introducción a través de C. I. Lewis en el contexto de la

con este par de experiencias, también pasa con las experiencias del azul y del verde; y con las experiencias de cualquier otro color, desde luego. Cada una de ellas cuenta con un *quale* particular, con un *quale* distintivo que permite al sujeto identificar en qué experiencia se encuentra. ¿Cabe pensar acaso en la experiencia de un color sin *quale* alguno? El dolor, el hambre y la sed tienen, asimismo, *qualia* (plural de 'quale') o qué-se-siente particulares – *qualia* sumamente desagradables si se me permite decirlo.

También exhiben *qualia* las experiencias olfativas (piénsese en el olor de las rosas o en el de un cabello quemándose), las experiencias gustativas (considérese la diferencia entre lo dulce y lo amargo, así como la diferencia entre lo ácido y lo salado), las experiencias táctiles (compárese la experiencia de lo rugoso con la experiencia de lo liso, así como la experiencia de lo seco con la experiencia de la húmedo), las experiencias auditivas (tráigase a la memoria el sonido de un conjunto de trompetas y el sonido de las olas del mar muriendo lentamente sobre la arena), las emociones (recuérdese los qué-se-siente del amor y del odio, de la envidia y el orgullo), los humores (piénsese en la calma y el aburrimiento) y las sensaciones corporales (el frío, el calor, la picazón, los hincos, el ardor, el hormigueo, las cosquillas y el orgasmo, entre otras). Consecuentemente, hablamos de los *qualia* en referencia a rasgos fenomenológicos cualitativos y no intencionales de ciertos estados mentales. Cabe hacer dos aclaraciones al respecto.

En primer lugar, los *qualia* como propiedades solo son atribuibles a estados mentales que sean conscientes. Si como consideran algunos, hay estados mentales inconscientes, estos no tienen *qualia*. El reino de lo inconsciente, si lo hay, no presenta aspectos

discusión sobre los *sense data*, véase Keeley (2009).

cualitativos de los que el sujeto esté al tanto. ¿Cómo podría dicho sujeto estar al tanto (o percatarse) de estos aspectos si se encuentran en la esfera de lo inconsciente y son, por tanto, ellos mismos inconscientes? No hay *qualia* sino hay conciencia de ellos. Es más, hablar de *qualia* inconscientes encierra una contradicción. La posibilidad de que haya *qualia* de los cuales no se tenga conciencia está excluida *a priori*. No tiene sentido afirmar que yo experimente el *quale* de cierto estado mental sin que me dé cuenta de ello. Un *quale* supone un darse cuenta (una percatación).

En este punto, algún sutil polemista podría decir que no es necesario que los *qualia* de estados mentales inconscientes también sean inconscientes. Según esto, el estado mental es inconsciente, pero su *quale* puede ser consciente. Esta distinción es falaz. Si el *quale* atribuido a cierto estado mental es consciente, ese estado mental también debe ser consciente. Pues un *quale* se asocia directamente a un estado mental determinado. Si soy consciente del *quale*, soy consciente del estado mental del cual es propiedad. Considérese que es consustancial al *quale* el permitir la identificación de cierto estado mental. No puede haber *qualia* que no asociemos con estados mentales. No puede haber *qualia* puros. El polemista podría incurrir en más sutilezas y argumentar en favor de que los *qualia* puros sí son posibles para, así, conferir de manera indirecta crédito a la posibilidad de que haya *qualia* de estados mentales inconscientes. Aduciría en esta línea que yo puedo tener el *quale* de la experiencia auditiva del batir de un conjunto de tambores sin tener ahora mismo esa experiencia auditiva, sin tener ningún tambor frente a mí. Puedo tener el *quale* simplemente como parte de un recuerdo o como producto de la imaginación. Concedemos esta posibilidad, pues, por ejemplo, también es concebible que tenga el *quale* del oír barritar a un elefante sin oír barritar a ningún elefante ahora mismo. Pero afirmamos que de esta posibi-

lidad no se sigue la afirmación de *qualia* puros, pues los *qualia* auditivos en cuestión, sin ser propiedades de experiencias auditivas, son propiedades de otros estados mentales, a saber, propiedades del recuerdo o de la imaginación e incluso de la alucinación. Los *qualia* no pueden dejar de ser propiedades. La idea de *qualia* puros no es concebible, como tampoco lo es la de *qualia* inconscientes o *qualia* conscientes de estados mentales inconscientes.

En segundo lugar, no está claro que todos los estados mentales conscientes presenten *qualia*. Expliquémonos, siguiendo la tradición de Brentano (1874), ha sido sólito afirmar que el rasgo específico de lo mental sea la intencionalidad, es decir, el representar algo. Los estados mentales no van por ellos mismos. Van por otras cosas. Representan otras cosas a diferencia de lo que ocurre con lo físico. La imagen mental de un árbol tiene intencionalidad. Representa a ese árbol. Está dirigida a ese árbol. La caída de una hoja de ese mismo árbol, por lo contrario, no representa nada. Es un hecho del mundo y va por sí mismo. Pero, siendo precisos, y *pace* Brentano, no obstante ser la intencionalidad un rasgo específico o, si se quiere, exclusivo de lo mental, no es un rasgo definitorio de lo mental. Es cierto que todo aquello que presente contenido intencional sea un estado mental, pero no es cierto que todos los estados mentales presenten contenido intencional. En otras palabras, la intencionalidad es condición suficiente de lo mental, pero no condición necesaria. Puede haber estados mentales que carezcan de ella y presenten únicamente contenido fenoménico, esto es, un *quale*. Intuitivamente, la experiencia de haberse golpeado el nervio del codo no tiene contenido intencional, es decir, no representa nada. Pero sí tiene un *quale* característico si no se sabe cuál es, hágase el experimento debido. Lo mismo ocurre con las experiencias de escalofríos y de recibir un pisotón, así como con todo aquello que a grandes rasgos pueda considerarse como sensacio-

nes. De este modo, encontramos una distinción entre géneros de estados mentales. Están, por un lado, aquellos cuyo contenido es puramente fenoménico o cualitativo, aquellos estados mentales que al decir de G. E. Moore (1903) no están conceptualizados y a los cuales llama “sensaciones.” Y están, por otro lado, los estados mentales que tienen contenido intencional. Moore llama “pensamientos” a estos últimos. Los estados mentales en referencia son todos aquellos que representen algo. Lo que no queda claro es si estos pensamientos presentan *qualia*.

¿Hay *qualia* que sean propiedades de pensamientos, digamos, propiedades de creencias o de dudas? La respuesta puede ser afirmativa si se considera a la certeza como un aspecto subjetivo, como algo que se siente. Pues, efectivamente, si dudo acerca de la verdad de cierta proposición (que hay más de un millón de habitantes en Oslo, por ejemplo) es porque no siento el mismo grado de certeza respecto de ella que el que siento en relación con una proposición en la cual creo (que Fidel Castro es cubano, por ejemplo). Es más, hay creencias respecto de las cuales albergo mayor certidumbre que por otras. Intuitivamente, a pesar de creer ambas cosas, que los cuerpos sean extensos y que Luis Castañeda Lossio sea el actual alcalde de Lima, puedo decir que creo más en lo primero que en lo segundo, pues el grado de certeza que experimento en relación con la primera creencia es mayor que el que experimento en relación con la segunda. Mi seguridad o certeza varía en función del contenido intencional de la creencia. Podría trazar un continuo que vaya desde aquellas actitudes proposicionales en cuyo contenido intencional no crea nada (que no existo, por ejemplo) hasta aquellas otras respecto de cuyo contenido intencional experimento absoluta certeza (que existo, por ejemplo). El centro de este continuo estaría constituido por las actitudes proposicionales en relación con cuyo contenido intencional alber-

go duda (que voy a dejar de existir en cinco años, por ejemplo). Todo este continuo estaría marcado por la intensidad de mi certidumbre, por cuán seguro me sienta de la verdad de las proposiciones en cuestión. Si esto es así, podría hablarse de la intensidad de la creencia como se habla de la intensidad de la picazón. Esta creencia es mayor que esta otra, así como esta picazón es mayor que esta otra.

¿Pero hay algo así como un *quale* característico de la creencia y como sucede con la picazón, que tiene un *quale* particular? Preguntamos si hay algún rasgo fenomenológico cualitativo y no intencional que permita decir “esto es una creencia y esto otro no” (“esto se siente como una creencia y esto otro no”), independientemente del grado de seguridad o certeza que tengamos, como ocurre cuando decimos “esto es una picazón y esto otro no” (“esto se siente como una picazón y esto otro no”), independientemente de la intensidad que se tenga o como dijo frente a la prensa en una memorable y aguda distinción fenomenológica Luis Enrique Martínez, director técnico del Fútbol Club Barcelona, tras sufrir dos derrotas consecutivas con su equipo entre octubre y noviembre de 2014: “No estoy preocupado... estoy cabreado” (La Liga, 2014, 01:54).

Alguien podría responder que la *sensación* de seguridad en sí misma es el *quale* de la creencia, pues no hay otro estado mental por el que experimentemos seguridad. O sea, solo se experimenta seguridad en relación con creencias y, por tanto, la seguridad es el *quale* de las creencias en tanto estados mentales conscientes. Respondiendo de esta manera el defensor de la tesis de que las creencias tienen *qualia* distintivos habrá hecho flaco favor a su posición. Pues de lo que ha dicho se sigue que no es la creencia, en tanto estado mental, la que tiene por propiedad el *quale* de la seguridad, sino que la creencia, en tanto estado mental, *causa* la

sensación de seguridad y esta sensación de seguridad es la que tiene por propiedad ese *quale* tan característico, el *quale* de la experiencia de seguridad. Brevemente dicho, la seguridad es una sensación, no un *quale*. Es un estado mental, no la propiedad de un estado mental. En consecuencia, no distinguimos entre creencias y dudas porque tengan *qualia* distintivos, sino por los *qualia* de las sensaciones que causan. Así como una creencia causa sensación de seguridad o certeza en mayor o menor grado, también puede causar sensación de indignación (que he sido detenido sin razones por la policía, por ejemplo), de alivio (que mi madre ha superado con éxito su operación al corazón, por ejemplo) o de entusiasmo (que Perú ha clasificado a Rusia 2018, por ejemplo).

Si se cuestiona que las creencias en tanto tipo de estados mentales tengan *qualia*, también habrá que cuestionar que las creencias en tanto casos o episodios tengan *qualia*. En efecto, no encontramos mayores razones para pensar que haya *qualia* que nos permitan distinguir una creencia de otra, entre las muchas —y potencialmente infinitas— que se puede tener. Siendo específicos, ¿hay *qualia* que me permitan distinguir entre la creencia de que Lima es la capital de Perú y la creencia de que Guayaquil es una ciudad de Ecuador?, ¿hay acaso qué-se-siente diferentes entre una creencia y la otra? ¿Qué se siente creer que $2 + 2$ es igual a 4? Los episodios de creencias se diferencian en función de su contenido intencional y no de su contenido fenoménico, ni siquiera es seguro que tengan contenido fenoménico.

Ahora bien, tal como se ha hablado de los *qualia* hasta aquí, esto es, a modo de aspectos no intencionales de estados mentales eminentemente cualitativos como pueden ser las experiencias perceptuales, las emociones, los humores y las sensaciones corporales, resulta incontrovertible la afirmación de que hay *qualia*. Los hay. El escepticismo en relación con ellos sería suicida y hasta

motivo de hilaridad. No hay duda de que los *qualia* marcan buena parte de nuestra experiencia consciente. Como precisa Crane (2001), los *qualia* son “propiedades no intencionales cuya instanciación explica (o parcialmente explica) el carácter fenoménico de los estados cualitativos”² (p. 76). El escepticismo en relación con los *qualia*, que tiene como autor paradigmático a Dennett (1988), solo puede entenderse considerando que el blanco de su crítica no es el discurso sobre *qualia* empleado en los párrafos anteriores, sino el que se refiere a estas propiedades caracterizándolas como intrínsecas, inefables, infalibles, atómicas, privadas, directas e irreducibles a lo físico³. Cabría dudar o de la realidad efectiva de los *qualia* o de la suficiencia del materialismo para explicarlos si sus características son estas. Repasémoslas.

Atribuir condición de *intrínsecos* a los *qualia* equivale a considerarlos como propiedades monádicas de los estados mentales. Esto quiere decir que los *qualia* no son relaciones. Tal cuadro nos invita a pensar que son propiedades que los estados mentales tienen por sí mismos y que no es correcto pretender entenderlas por las relaciones que el estado mental del caso tenga con otros estados mentales, *inputs* sensoriales u *outputs* conductuales; es decir, no sería correcto entender los *qualia* según el esquema funcionalista de caracterización de los estados mentales. Si los *qualia* son intrínsecos, no hay modo de entenderlos en términos relacionales ni causales. Si los *qualia* son intrínsecos, no se les puede funcionalizar.

Se suele creer, asimismo, que los *qualia* son *inefables*. ¿Hay, piénseselo bien, algún modo de verbalizar el *quale* de mi experiencia perceptual del rojo? Poetas y científicos quedan cortos frente a este reto. Nada puedo decir en relación con ese *quale*, salvo que

2 Mi traducción.

3 Se recomienda sobre el tema a Orlando (1995) y Pérez (2002).

es algo que siento como parte de mi experiencia visual. Pero dar cuenta del *quale* del rojo diciendo que es lo que se siente al tener —recordar o alucinar— la experiencia visual de, por ejemplo, el color de la sangre es no explicar nada e incurrir en un círculo vicioso. Los *qualia* son algo que no se puede comunicar en términos verbales. Simplemente se sienten. No se les puede describir. Son ininteligibles a menos que se les haya sentido. Si se les tiene —o se les ha tenido— se entiende qué son; y si jamás se les tuvo, no hay manera de entender qué son. Quien no haya experimentado el *quale* del rojo no podrá entender en qué consiste hasta que él mismo lo experimente. Cualquier intento de descripción o explicación está condenado al fracaso, pues, digamos, nunca podrá hacer entender qué es el *quale* del rojo a un invidente de nacimiento o a una persona que padezca de acromatopsia (también de nacimiento). Alguien dirá que el artista cibernético Neil Harbisson, sufriendo de acromatopsia, es decir, no siendo capaz de ver los colores cromáticos, y estando limitado a la visión del blanco, el negro y las diferentes escalas de gris, tiene no obstante —y, en consecuencia, logra entender qué es— el *quale* del rojo gracias a la información que recibe de parte de su “ojo artificial”⁴. El *quale* del rojo no sería, por tanto, inefable. En efecto, Harbisson identifica con éxito el color rojo gracias al sensor que tiene adherido a su cuerpo. Repárese, sin embargo, en el hecho de que este aparato no le concede una experiencia visual del rojo, sino, más bien, una experiencia auditiva del mismo al convertirla frecuencia luminosa de la longitud de onda correspondiente al color rojo en frecuencia sonora y transmitírsela vía ósea. Harbisson no ve colores. Los es-

4 Cf. Harbisson (2012) y La Ciudad de las Ideas (2013). Agradezco a Harbisson y a su equipo de colaboradores por el envío de su manuscrito no publicado, que puede encontrarse en la bibliografía como Harbisson (2016). Este material me ayudó a entender mejor su caso.

cucha. Él ve la sangre y escucha su color. Su sensor no le explica ni le concede la vivencia del *quale* visual del rojo. Este *quale*, al igual que el *quale* auditivo del rojo, continúa siendo inefable.

También se dice de los *qualia* que son *infalibles*. Entiéndase esto como la afirmación de que nadie se equivoca al señalar que siente tal *quale* o tal otro. Las razones que suele aducirse suenan convincentes. Si digo que me duele la muela, ni siquiera un calificado odontólogo —y tampoco un psiquiatra— me puede corregir. ¿Y qué médico o neurofisiólogo, asimismo, podría convencerme de que en realidad no tengo los *qualia* del hambre ni de la sed cuando digo y pienso que tengo hambre o que tengo sed? De igual modo, aunque no tenga una manzana roja frente a mí, y se trate de un objeto de mi alucinación, nadie puede decirme que no estoy teniendo el *quale* del rojo. Nadie me puede corregir. Somos, de acuerdo con esto, inmunes al error en lo concerniente a la vivencia de *qualia*. Si se cree que se les siente, no hay duda de que se les siente. No importa que su causa no sea real y que se trate de algún tipo de alucinación. Puede, en este sentido, recordarse la etiología de casos como aquellos denominados “dolores de miembros fantasmas”⁵. Hay un episodio de esta naturaleza cuando, por mencionar un ejemplo, cierto individuo cuya pierna derecha haya sido amputada afirma que le duele el talón del pie derecho. El dolor es real. El sujeto lo siente, aunque carezca del miembro. No hay equivocación de su parte.

Por otro lado, se dice de los *qualia* que son *atómicos* y, por tanto, *inanalizables*. No se trata de propiedades complejas que puedan ser descompuestas. No se nos presentan por partes o fragmentadas. Las vivenciamos como un todo. Son unidades simples. Esta caracterización está relacionada a la consideración de los *qualia*

5 Cf. Nikolajsen & Jensen (2001), Ramachandran & Hirstein (1998) y Halligan (2002).

como intrínsecos e inefables. También guarda vínculo con la caracterización de los *qualia* como privados.

Los *qualia* son *privados* porque solo son accesibles al sujeto que los vive. Este acceso se daría exclusivamente a través de la introspección. Así, no es posible la observación de *qualia* desde el punto de vista de una tercera persona. No se puede ver a los *qualia* desde el exterior del sujeto que los vive. La observación está restringida a la perspectiva de la primera persona, esto es, a la perspectiva del yo. Nadie puede observar mi *quale* de rojo. Podemos hablar del qué-se-siente de los estados mentales solo en relación con nuestro propio caso. Nagel (1974) da cuenta de este carácter privado de los *qualia* cuando argumenta que; no obstante, la posibilidad de contar con una descripción física conspicua de los estados de un murciélago, nunca podremos saber qué se siente ser un murciélago a menos que seamos un murciélago, esto es, a menos que nuestra observación se realice desde la perspectiva de la primera persona. Tampoco nadie podrá saber qué se siente mi *quale* de rojo a menos que sea yo mismo.

La introspección recién mencionada no nos ofrece un conocimiento inferencial o mediato. La conciencia de los *qualia* resulta directa. Por eso se dice de los *qualia* que son *directos*. No son consecuencia de una inferencia.

Ahora bien, la característica más controvertida de los *qualia* para la filosofía es la de su presunta *irreductibilidad* en relación con lo físico. Esta característica guarda estrecha relación con la condición privada y con la condición inefable de los *qualia*. Si únicamente quien vive el *quale* puede saber qué es y que efectivamente se da, y no importa, por tanto, cuán buena sea la descripción física que un tercero haya hecho del organismo que lo viva, entonces los *qualia* no guardan relación de identidad con los estados físicos. Si esto es así, lo mental no se reduce a lo físico. No se puede obtener una

comprensión de lo primero a partir de una comprensión de lo segundo. Esta es la conclusión del llamado “argumento del conocimiento” de Jackson (1982). El *Gedankenexperiment* sobre el cual se funda dicha conclusión propone imaginar el caso de Mary, una brillante neurofisióloga especializada en el fenómeno de la visión, quien, por motivos que ignoramos, ha vivido recluida desde siempre en una habitación donde solo puede tener experiencia del blanco y del negro⁶. Todas sus investigaciones fueron llevadas a término desde este recinto, apoyándose de un televisor cuando era necesario, este televisor también era a blanco y negro. Mary sabe todo lo físico que hay que saber cuando los humanos ven tomates maduros o el cielo despejado. Descubre la longitud de onda correspondiente a cada color, cómo estimula nuestras retinas y el proceso nervioso que nos hace finalmente decir “El cielo es azul” o “Los tomates son rojos”. Nada físico se le escapa del color ni de nuestras experiencias visuales de color. Imagínese ahora que Mary abandona la habitación donde vivía encerrada. Su experiencia visual del mundo ya no será a blanco y negro. Según Jackson, resulta obvio que Mary va a aprender algo nuevo sobre los colores cuando salga de su encierro. Va a aprender en qué consisten sus *qualia*, una cuestión sobre la cual no tenía la más mínima idea, a pesar de saber todo lo físico que puede saberse acerca de los colores⁷. Esto constituye una razón

6 Lo suyo, a diferencia de la acromatopsia congénita de Neil Harbisson, es una suerte de acromatopsia de hecho inducida por las condiciones ambientales del espacio donde se encuentra encerrada.

7 Apreciamos una situación similar en el episodio “La medida de un hombre” de la serie televisiva *Viaje a las estrellas: La nueva generación* (Snodgrass & Scheerer, 1989). En este, el agente Data, un androide inteligente, se rehúsa a ser desmantelado, pues aunque al ensamblársele de nuevo contará con todas sus memorias, se habrá perdido irremediablemente los *qualia* que ahora mismo él asigna a estas memorias. Tales pasarán a ser mera información y dejarán de ser experiencias vividas. Ya no tendrán significado emotivo para él. La pura descripción física o funcional de sus

poderosa en favor de la tesis de que los *qualia* son irreductibles a lo físico. La idea es que se puede dar una descripción física —y también funcional— completa del mundo y no decir todo lo que hay. La conciencia fenoménica se resiste a ser capturada. Mostremos a continuación un breve ejemplo de ello con el problema de los *qualia* ausentes para el funcionalismo.

2. *Qualia* ausentes y funcionalismo

NedBlock (1978) populariza el argumento de los *qualia* ausentes a modo de crítica al carácter completo de la explicación funcionalista de la naturaleza de los estados mentales. Sabemos que, de acuerdo con esta concepción, los estados en los cuales se encuentra nuestra mente son idénticos a estados funcionales de cierta tabla de máquina. Es decir, la naturaleza de nuestros estados mentales es el rol causal específico que desempeña cada uno de ellos dentro del sistema de lo mental. Esto supone que toda organización con estados mentales es describable mediante una tabla de máquina en particular. Digamos que los *inputs* y *outputs* se especifican como impulsos neuronales. Ahora imaginemos un robot que a nivel exterior muestre extrema similitud con el cuerpo de un ser humano, pero cuyo interior presente una historia por completo diferente. Su figura es humana. Su piel luce como humana. Su conducta se asemeja a la de un humano. Grita, camina, los cabellos se le erizan. Su cabeza, no obstante, carece de cerebro. Está poblada de pequeños hombres, de homúnculos. En esta cabeza hueca, las “neuronas” asociadas a los órganos sensoriales se conectan con una hilera de luces. Estas luces especifican los *inputs* posibles del sistema. Sus *outputs*, mientras tanto, están determinados por un conjunto de botones que se

estados no capturará el aspecto cualitativo de sus recuerdos. La tesis subyacente es que los *qualia* son irreductibles a lo físico y también inefables.

conectan con las “neuronas” motoras. Los estados singulares en los que se encuentra el robot son marcados por un letrero que cuelga dentro de la cavidad “craneal.” Cada hombrechito está asociado a un letrero específico. La tarea del conjunto es implementar una tabla de máquina idéntica a la de una mente humana. Así, en caso de que el letrero colgado diga “Alivio” (‘G’ en la figura 1) y se encienda la luz correspondiente a una piedra cayendo sobre el pie del robot (la luz del *input* sensorial I_{17} en la figura 1), el hombrechito encargado del letrero “Alivio” (el G-man de la figura 1) debe cambiarlo por el de “Dolor” (‘M’ en la figura 15) y presionar el botón que haga exclamar “¡Au!” al robot (botón del *output* conductual 191 en la figura 1). El estado en el que ahora se encuentra este robot tiene la misma descripción funcional que nuestro estado de dolor. Si esto es así, el funcionalista nos permitirá decir que el robot se encuentra en estado de dolor, que el robot tiene dolor. Pero, ¿es este realmente el caso? ¿Acaso ocurre que el robot *sienta* dolor?

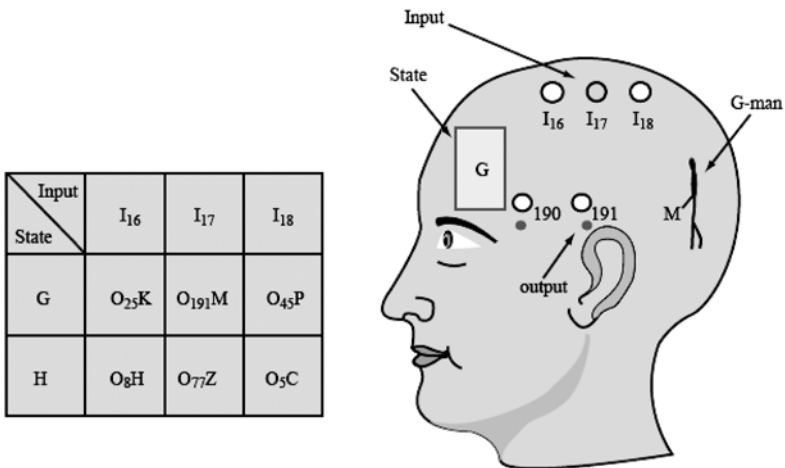


Figura 15. Robot de cabeza homuncular⁸.

8 Ilustración tomada de Byrne (2011).

La descripción funcional de nuestro estado puede ser idéntica a la del estado en el cual se encuentra el robot; sin embargo, hay algo en nuestro estado de dolor que *prima facie* no presenta el estado en el cual se halla el robot: el aspecto cualitativo. Nuestro estado de dolor presenta un *quale*; y el del robot de cabeza homuncular, no. Para nosotros, el encontrarse en estado de dolor se siente de cierto modo (tiene un *quale* particular); y parece ser que el robot encontrándose a ojos funcionalistas en estado de dolor no siente nada —no tiene *quale*— a pesar de los quizá desgarradores gritos que profiera a viva voz. Tal es el argumento de los *qualia* ausentes en contra del funcionalismo.

La ausencia de *qualia* en un sistema funcionalmente idéntico a otro que sí presenta *qualia* constituye un serio problema a la viabilidad del funcionalismo como teoría completa acerca de la naturaleza de los estados mentales. En el mejor de los casos, si la objeción es correcta, el funcionalismo tendría que ser considerado únicamente como una teoría parcial. No explicaría la naturaleza de todos los estados mentales, al menos no la de aquellos que presenten *qualia*. Pues, dada la posibilidad de un escenario como el recién descrito, el funcionalismo estaría obligado a afirmar que el robot de cabeza homuncular tiene dolor, ya que se encuentra en un estado cuya descripción funcional es idéntica a aquella que presentamos nosotros cuando nos encontramos en estado de dolor, cuando tenemos —y, por supuesto, sentimos/sufrimos— dolor. Sin embargo, esta atribución de dolor al robot parece ser un sin sentido. Consideramos que el robot no puede encontrarse en estado de dolor porque no puede sentir el *quale* del dolor. Consecuentemente, la descripción correcta del dolor no se ve agotada por una descripción funcional. El dolor, además de un rol causal, que es al final de cuentas todo lo que el funcionalismo puede advertir, es algo que se siente y el esquema de explicación funcionalista resulta nulo para la tarea de

captar este rasgo fenomenológico no intencional. El problema no es pequeño. La región de lo mental que no explicaría con suficiencia el funcionalismo es considerablemente grande, tanto en número como en importancia, ya hemos visto que los estados mentales con *qualia* incluyen experiencias perceptuales, sensaciones corporales, emociones, humores, entre otros.

El escenario del robot con cabeza homuncular puede sonar poco convincente a más de uno. Para estos escépticos, se trataría de una ficción filosófica malintencionada —y exagerada— propuesta con el fin deliberado de minar al tan popular y extendido funcionalismo. En tal sentido, estos mismos escépticos dirían que un organismo (robot) como el descrito implementa una mente humana como la nuestra únicamente en caso de que cuente con literalmente miles de millones de pequeños hombrecitos en su cabeza hueca que efectúen los cientos de miles de cambios posibles de tarjeta (cambios de estados) y presionen los miles de millones de botones de *outputs* conductuales posibles según, desde luego, cuál de los miles de millones de *inputs* sensoriales posibles se haya recibido. No cabe pensar que algo como esto, no digamos suceda, sino que sea simplemente posible, por lo menos en principio. Se requeriría de demasiados elementos y de una coordinación milimétrica entre ellos y, es más, de que todo ello ocurra dentro de una cabeza de forma y proporciones humanas. Block (1978) hace frente a estos escépticos. El argumento de los *qualia* ausentes es correcto más allá de la viabilidad o no de que un conjunto de homúnculos dentro de la cabeza de un robot con forma humana implementen la tabla de máquina correspondiente a nuestra mente. Si un experimento mental como este ha parecido exagerado, entonces ya no pensemos en homúnculos. Pensemos ahora en los miles de millones de habitantes de China.

Gracias a la intervención de sus gobernantes (recientemente convertidos al funcionalismo y no se pregunte cómo), todos estos individuos implementarán la tabla de máquina atribuible desde un punto de vista funcionalista a una mente humana y pondrán en movimiento el cuerpo del robot de marras durante el lapso de exactamente una hora. Ya no hay tarjetas colgadas con los nombres del estado en el cual se encuentra el sistema. Ahora el nombre del estado se muestra en el cielo mediante señales proyectadas por satélites, de modo que cualquiera de los miles de millones de chinos pueda notarlo y, en particular, aquellos chinos que deben poner manos a la obra cuando el robot entre en tal o cual estado particular. Además, se proporciona un radiotransmisor a cada uno de los habitantes de este país. Los radiotransmisores reciben las señales de los *inputs* sensoriales que afectan al robot. Ya que son miles de millones de *inputs* posibles, hay un *input* determinado para cada uno de los miles de millones de habitantes de China. Quien note el nombre de su estado particular proyectado en el cielo y reciba mediante radiotransmisor la señal de que su *input* sensorial se ha dado, presionará el botón del radiotransmisor que corresponda al *output* conductual del caso específico. Siguiendo con el ejemplo de párrafos más arriba, si al robot le cayó una piedra en el pie y estaba proyectada la señal de “Alivio” en el cielo, entonces un chino específico al cual le llega el aviso por radio tiene que presionar el botón que haga exclamar “¡Au!” al robot. Luego, el satélite proyectará la señal de “Dolor” en el amplio cielo del país-continente asiático.

Este sistema es posible. ¿Qué diríamos respecto de sus rasgos mentales? ¿Acaso los tiene? Los miles de millones de habitantes de China implementan la tabla de máquina correspondiente a la mente humana y, consecuencia de ello, el robot manifiesta conducta típicamente humana en respuesta a una serie de estímulos

sensoriales. Exclama “¡Au!” cuando le cae una piedra en el pie, por ejemplo. La pregunta es si el robot tiene experiencia consciente. ¿Se siente de algún modo ser ese robot, así como Thomas Nagel (1974) sugiere que se siente de algún modo [*whatitislike*] ser un murciélago? ¿Siente algo este androide cuando le cae una piedra sobre la parte de su cuerpo que llamaríamos pie y termina exclamando “¡Au!”? Si se dice que el cuerpo del robot no siente nada, pues simplemente es un montón de acero y hojalata que recibe señales e instrucciones vía radio, ¿cabría, entonces, pensar que sí hay experiencia consciente —que sí hay *qualia*— en su cerebro, que en este caso es la nación china en su totalidad? Intuitivamente, no hay *quale* de dolor por ningún lado. A pesar de que China pueda hacer las veces de un cerebro que implemente la tabla de máquina correspondiente a la mente humana, no se nos ocurre pensar que algo así como una conciencia individual emane de la totalidad de habitantes chinos, que China como un todo sienta el dolor. No obstante esto, la descripción funcional de los estados que implementan los miles de millones de chinos es idéntica a la descripción funcional de nuestros estados mentales. En nosotros hay conciencia fenoménica, mientras que en China no hay un solo atisbo de ella. Esta reformulación del experimento mental del robot con cabeza homuncular —con cerebro chino, diríase ahora— muestra que el funcionalismo debe admitir como posible un escenario de *qualia* ausentes. Y si esto es así, debe admitir de sí mismo que es una teoría parcial y no completa sobre lo mental.

Otro escenario de *qualia* ausentes que ha sido bastante comentado en las últimas dos décadas es el propuesto por David Chalmers (1996) en relación con la posibilidad —al menos metafísica— de *zombies*. Planteemos el asunto. No está en discusión que dispongamos de la capacidad de concebir. La empleamos de

continuo en nuestra vida diaria y en momentos un tanto más reflexivos, sean estos filosóficos o científicos, ideando argumentos de sillón así como experimentos mentales. Gracias a esta capacidad, representamos diversos arreglos del mundo mediante palabras, conceptos o imágenes. Ahora bien, una creencia metafísica profundamente arraigada en nuestro pensamiento consiste en señalar que lo concebible es también algo posible. No está claro qué justifica este salto, el ir de la afirmación de que algo es concebible a la conclusión de que ese mismo algo sea posible. Resulta, sin embargo, manifiesto que da lugar a tesis en lo absoluto triviales. Recuérdese, por ejemplo, que Descartes (1641) funda su célebre distinción entre mente y cuerpo asido de este principio de inferencia. En sus *Meditaciones*, sostiene que, dado que la mente es concebible aparte del cuerpo, es, en consecuencia, posible que la mente esté aparte del cuerpo, y, así, que se trata de sustancias distintas. En años más recientes, el salto de lo concebible a lo posible ha sido usado no tanto para abogar en favor de una especie de dualismo sustancial, sino, más bien, para poner en cuestión a las concepciones fisicalista y funcionalista en relación con el problema mente-cuerpo. Tenemos, en este sentido, el llamado argumento de los *zombies*. De acuerdo con este, los *zombies* son concebibles y, por lo tanto, posibles. Tal conclusión iría directamente en contra de cualquier visión materialista de lo mental.

El concepto de *zombie* hace referencia a una réplica exacta de un humano a nivel físico y funcional en determinado momento. Los *zombies* son duplicados. Piénsese en un individuo humano cualquiera. Por ejemplo, tómese el caso del político César Acuña Peralta, líder y fundador del Partido Alianza para el Progreso (APP). Un *zombie* de César Acuña es un organismo que a nivel físico y funcional constituye una réplica de César Acuña. Esto es, todos los estados físicos que tenga César Acuña, y únicamente

esos estados físicos, también estarán presentes en el mismo momento en el *zombie* singular en cuestión. Los dos son individuos de carne, sangre y hueso dispuestos en el mismo orden. Por lo mismo, y considerándose la superveniencia de los estados funcionales en relación con los estados físicos, todos los estados funcionalmente descriptibles que presente César Acuña también se hallan presentes en el mismo *zombie* y en el mismo momento.

César Acuña y su *zombie* lucen igual y se comportan igual. Reaccionan a los mismos estímulos de la misma manera. Física y funcionalmente son los mismos. La única diferencia es que el *zombie* de César Acuña no tiene conciencia, ¡pequeño detalle!, por cierto. César Acuña *siente* dolor y exclama “¡Au!” cuando una piedra cae sobre su pie. El *zombie* exclama “¡Au!” cuando una piedra cae sobre su pie, pero no siente dolor. Ser César Acuña, además, se siente de alguna manera. Tiene un *quale* distintivo que el político César Acuña —y solo él— debe de experimentar. Sin embargo, ser el *zombie* de César Acuña no se siente de modo alguno. No hay un *quale* distintivo. El *zombie* no tiene ninguna experiencia consciente. Las luces están encendidas, pero no hay nadie en casa. El interior del *zombie* es un recinto experiencialmente vacío. El *zombie* no siente nada. No tiene conciencia fenoménica de nada. Esta situación coloca en serios aprietos tanto a los fisicalistas como a los funcionalistas.

En primer lugar, si los estados mentales son estados físicos, el *zombie* tendría que presentar exactamente los mismos estados mentales atribuibles a César Acuña ahora, pues es una réplica física exacta de César Acuña ahora. Pero César Acuña siente dolor cuando una piedra cae sobre su pie, mientras que el *zombie* no siente nada cuando una piedra cae sobre su propio pie. Entonces, está claro que el *zombie* no tiene los mismos estados mentales, al menos no los mismos estados mentales con *qualia* que cuente con

estados mentales puramente intencionales o representacionales, si los hay, resulta indiferente al argumento. Si esto es así, la tesis fisicalista es falsa.

En segundo lugar, si los estados mentales son estados funcionales, el *zombie* tendría que presentar los mismos estados mentales que César Acuña. Pero este no es el caso. Ya se ha dicho que César Acuña siente dolor cuando una piedra cae sobre su pie, mientras que el *zombie* no siente nada. Si esto es así, la tesis funcionalista es falsa.

Tenemos, entonces, otro ejemplo de dos sistemas con las mismas descripciones funcionales, no obstante el hecho de que uno de ellos presente *qualia* y de que en el otro no haya *qualia*. Si los estados mentales son estados funcionales, el filósofo y científico funcionalista deberá afirmar que el *zombie* tiene dolor, pues presenta el estado funcional que en el caso de César Acuña se reconoce como dolor. Obviamente, decir que hay estados mentales que son dolores aunque no presenten *qualia* resultaría suicida para la posición funcionalista en tanto pretende ser un modelo completo de la naturaleza de los estados mentales. Se estaría echando por la borda a los *qualia*, a pesar de que para nosotros no quepa pensar en dolores sin *qualia*. Esto haría del funcionalismo un modelo cuestionable de lo mental.

Terminemos ofreciendo brevemente un ejemplo más de ausencia de *qualia* en un sistema cuya descripción funcional sea idéntica a la de un sistema que sí tiene *qualia*, ya que el escepticismo inicial sobre la viabilidad del robot de cabeza homuncular puede presentarse nuevamente, ahora en relación con la viabilidad de los *zombies*. Esta vez apelamos a un *Gedankenexperiment* de Palmer (1999) que da cuenta de estados de cosas físicamente posibles aunque para refutar una tesis sea suficiente con un escenario meramente concebible como lo ha sido el de los *zombies*. Este autor describe minuciosamente el funcionamiento de una máquina que

detecta los colores con éxito. Esta máquina no solo es concebible, sino que como cuestión de hecho es realizable. Añádase el detalle de que entre ella y nosotros hay un isomorfismo causal. Esto es, los dos sistemas contamos con la misma descripción funcional. Procesamos la información a partir de la luz de la misma manera y nuestras respuestas a “¿qué color es ese?” son las mismas. Sin embargo, intuitivamente, la máquina no tiene experiencia del color, algo que nosotros sí tenemos. No importa que entre nosotros, los humanos, nuestros *qualia* estén invertidos. El punto es que tenemos *qualia*, cosa que no tiene la máquina detectora de colores. Este hecho, a nuestro parecer intuitivo, resulta imperceptible para el funcionalismo. Esta teoría es incapaz de efectuar una discriminación entre nuestras experiencias fenoménicas de color y la completa ausencia de experiencias de este tipo en la máquina. Para ella los dos nos encontramos en los mismos estados.

Bibliografía

- BLOCK, N. (1978/1980). Troubles with functionalism. En su *Readings in philosophy of psychology*, Vol. I (pp. 268-305). Cambridge, Massachusetts: Harvard University Press.
- BRENTANO, F. (1874). *Psychologie vom empirischen Standpunkte*, Leipzig: Duncker & Humblot.
- BYRNE, A. (2011). Functionalism [Material 8 del curso “Minds and machines”]. Departamento de Lingüística y Filosofía, Massachusetts Institute of Technology. Recuperado de: http://ocw.mit.edu/courses/linguistics-and-philosophy/24-09-minds-and-machines-fall-2011/study-materials/MIT24_09F11_functionalism.pdf [Consulta: 25 de enero de 2016].
- CHALMERS, D. (1996). *The conscious mind*. Nueva York: Oxford University Press.
- CRANE, T. (2001). *Elements of mind*. Oxford: Oxford University Press.

- DENNETT, D. C. (1988). Quining qualia. En A. Marcel & E. Bisiach (Eds.), *Consciousness and contemporary Science* (pp. 42-77). Oxford: Oxford University Press.
- DESCARTES, R. (1641/2010). *Meditaciones acerca de la filosofía primera*. Bogotá: Universidad Nacional de Colombia.
- HALLIGAN, P. W. (2002). Phantomlimbs: Thebody in mind. *CognitiveNeuropsychiatry*, 3(7), pp. 251-269.
- HARBISSON, N. (2012, Junio). Neil Harbisson: I listen to color [Archivo de video]. Recuperado de https://www.ted.com/talks/neil_harbisson_i_listen_to_color [Consulta: 8 de febrero de 2016].
- HARBISSON, N. (2016). *Neil Harbisson: A collection of essays written by the cyborg artist exploring the art of creating new senses and extending our perception*. Nueva York: Manuscrito de publicación pendiente.
- HILL, C. (1997). Imaginability, conceivability, possibility and the mind-body Problem. *Philosophical Studies*, 1(87), pp. 61-85.
- JACKSON, F. (1982). Epiphenomenal qualia. *Philosophical Quarterly*, 127(32), pp. 127-136.
- KEELEY, B. L. (2009). The early history of the *quale* and its relation to the senses. En J. Symons y P. Calvo (Eds.), *The Routledge companion to philosophy of psychology* (pp. 71-89). Londres: Routledge.
- LA CIUDAD DE LAS IDEAS [La Ciudad de las Ideas]. (2013). Neil Harbisson - La Ciudad de las Ideas 2012 "The Magic of If" [Archivo de video]. Recuperado de <https://www.youtube.com/watch?v=DYzFoP-gvmg>. Consulta: 8 de febrero de 2016.
- LA LIGA [LaLiga]. (2014). Luis Enrique: "No estoy preocupado porque la actitud ha sido buena. Estoy cabreado, como ellos" – HD [Archivo de video]. Recuperado de https://www.youtube.com/watch?v=d8WKKH_UEnQ. Consulta: 8 de febrero de 2016.
- LEVINE, J. (1998). Conceivability and the metaphysics of mind. *Nous*, 4(32), pp. 449-480.
- MOORE, G. E. (1903). The refutation of idealism. *Mind*, 12, pp. 433-453.
- NAGEL, T. (1974). What is it like to be a bat? *Philosophical Review*, 83, pp. 435-456.

- NAGEL, T. (1998). Conceiving the impossible and the mind-body problem. *Philosophy*, 73, pp. 337-352.
- NIKOLAJSEN, L. & Jensen, T. S. (2001). Phantom limb pain. *British Journal of Anaesthesia*, 1(87), pp. 107-116.
- ORLANDO, E. (1995). Contenido y conciencia: El debate en torno a los *qualia*. *Dianoia*, 43(43), pp. 1-29.
- PALMER, S. E. (1999). Color, consciousness, and the isomorphism constraint. *Behavioral and Brain Sciences*, 22, pp. 923-989.
- PÉREZ, D. (2002). Los *qualia* desde un punto de vista naturalista. *Azafea*, 4, pp. 65-83.
- RAMACHANDRAN, V. S. & Hirstein, W. (1998). The perception of phantom limbs. *Brain*, 121, pp. 1603-1630.
- SNODGRASS, M. M. (Guión), & Scheerer, R. (Director). (1989). The measure of a man [Episodio de serie de televisión]. En G. Roddenberry & R. Berman (Productores), *Star trek: The next generation*. Los Ángeles, California: Paramount.
- WORLEY, D. (2003). Conceivability, possibility, and physicalism. *Analysis*, 63, pp. 15-23.

Correspondencia:

David Villena Saldaña

Docente del Departamento Académico de Filosofía de la Facultad de Letras y Ciencias Humanas de la Universidad Nacional Mayor de San Marcos.

Correo electrónico: david.villena@cesfia.org.pe