

Mente y belleza, función sin forma y forma sin función

De ahí, además, que no quepa preguntarse si el cuerpo y el alma son una única realidad, como no cabe hacer tal pregunta acerca de la cera y la figura impresa en ella [...].

Aristóteles, *Acerca del alma* (II, 1, 421b)

Belleza es forma de la finalidad de un objeto en cuanto es percibida en él sin la representación de un fin.

Immanuel Kant, *Crítica del Juicio* (1790)

I. Introducción

Objetivos:

El objetivo de este trabajo es reflexionar acerca de los conceptos de forma y función de forma independiente, y abordar cada uno de estos conceptos desde ámbitos distintos de la filosofía. Así, el concepto de función se estudia en el contexto de la filosofía de la mente, por medio del estudio del funcionalismo, y el de forma en el contexto de la filosofía estética, partiendo del formalismo kantiano. Respecto a lo primero, si bien suele abordarse en la Filosofía de 1º de Bachillerato el debate acerca de si podría pensarse o no una máquina en los temas de antropología, no suele tomarse en consideración el argumento putnamiano de la realizabilidad múltiple que está en el origen del funcionalismo y que permite entender adecuadamente la idea de que los estados mentales son estados funcionales, y aquí aprovechamos la ocasión para profundizar en ello. En lo que respecta al formalismo kantiano, tanto en la Filosofía de 1º de Bachillerato como en la Historia de la Filosofía de 2º de Bachillerato, este suele tratarse solo desde el punto de vista de la filosofía moral, y apenas suele ahondarse en la *Crítica del juicio*, dejándose así de lado el pensamiento estético de Kant. No obstante, la distinción forma/contenido es fundamental en filosofía, y es importante que los alumnos sean capaces de verla no solo en el ámbito del conocimiento y de la moral, sino también en el de la experiencia estética y artística.

Contenidos:

La mente, función sin forma. Tras una explicación general del problema mente/cuerpo desde la perspectiva de la filosofía de la mente del último medio siglo, de las principales características de lo mental y las dificultades que conlleva, el estudio se centra en las teorías materialistas. Así, el trabajo se centra en el cambio de paradigma desde el reduccionismo, encarnado por la teoría de la identidad y el conductismo, al materialismo no reduccionista encarnado por el funcionalismo. Se recogen los principales argumentos de Hilary Putnam en contra del reduccionismo y se explica de forma pormenorizada (aunque sin entrar en sus distintas variedades) el funcionalismo. Por fin, se abordan los principales argumentos en contra de dicha teoría y se esbozan algunos de los contraargumentos de los filósofos funcionalistas.

La belleza, forma sin función. Aquí nos centramos en la *Crítica del juicio* kantiana. Se explica por una parte el juicio de gusto según Kant, esto es, “la facultad de juzgar lo bello”, un juicio que no proporciona conocimiento, es desinteresado y universal. Y a continuación se aborda su teoría de la belleza, que por ser estrictamente formal, nos permite ahondar precisamente en el concepto de forma, sobre todo mediante la distinción entre belleza libre y belleza adherente.

Materiales:

La reflexión en torno a los conceptos de función y de forma partirá de diversas fuentes. Por una parte, el estudio de la respuesta funcionalista al problema mente-cuerpo se abordará a partir de algunos fragmentos del artículo “La naturaleza de los estados mentales” de Hilary Putnam, pero sobre todo del relato de Daniel Dennett “¿Dónde estoy yo?”, sobre el que se plantean una serie de cuestiones específicas para evaluar los contenidos de filosofía de la mente. También, dentro de este apartado temático centrado en la filosofía de la mente, se ofrecen los principales argumentos con que se defienden muchas de las principales perspectivas acerca de lo mental, para que puedan ser formalizados o analizados y discutidos. Por fin, en el apartado de estética, se ofrecen varios textos de la *Crítica del juicio* de Kant para su comentario, así como una metodología del comentario de texto y una plantilla de corrección. También, se han seleccionado algunas imágenes que permiten comprender la diferencia entre belleza libre y adherente, y natural y artificial.

En lo que sigue, se aborda primero de forma teórica los contenidos, y a continuación, secuenciado por sesiones, la metodología específica para cada apartado, la forma de llevarlo al aula.

II. Contenidos teóricos

II.1. La mente, función sin forma

II.1.1. El problema mente-cuerpo en la filosofía contemporánea

En *Mentes, cerebros y ciencia*, John Searle caracteriza el problema mente-cuerpo de la siguiente forma:

[T]enemos una cierta representación de sentido común de nosotros mismos como seres humanos que es muy difícil casar con nuestra concepción ‘científica’ global del mundo físico. Nos pensamos a nosotros mismos como agentes conscientes, libres, cuidados, racionales en un mundo del que la ciencia nos dice que consta enteramente de partículas físicas carentes de mente y de significado. Ahora bien, ¿cómo podemos conjugar esas dos concepciones? ¿Cómo, por ejemplo, puede ser el caso de que el mundo no contenga otra cosa que partículas físicas inconscientes y que, con todo, contenga también conciencia? ¿Cómo puede un universo mecánico contener seres humanos intencionales -esto es, seres humanos que pueden representarse el mundo a sí mismos? ¿Cómo, para decirlo brevemente, puede un mundo esencialmente carente de significado contener significados? (Searle 2001, p. 17)

Todos entendemos aproximadamente qué es “lo físico”. Podemos partir de la definición cartesiana y entenderlo como aquello que posee extensión, o sencillamente como aquello que es corpóreo o que está compuesto por materia. No obstante, la materia ya no es lo que era y es difícil hablar de la extensión de las partículas subatómicas. Aun así, entendemos que hay una realidad dominada por las leyes de la física, que puede ser descrita con rigor y de forma puramente objetiva, que es razonablemente matematizable y predecible.¹ La ciencia, debido a su enorme éxito, parece abarcar todo lo real, de tal forma que existir materialmente y ser real aparecen como sinónimos.

¹ Por ello, aquí vamos a preferir el término clásico “materialismo” para referirnos al “fiscalismo”, en la lectura de estas páginas habrá que considerar ambos términos completamente sustituibles.

Pero el caso es que también hablamos de cosas tales como creencias y deseos, tenemos experiencia de aquello que llamamos "lo mental", y esto, a diferencia de lo físico, no parecemos entenderlo tan bien. Sabemos a qué se parece, porque lo vivimos, pero no resulta tan obvio que sea posible describir la mente con el rigor y la objetividad con que lo hacemos con respecto a las entidades físicas, no parece que la conciencia pueda ser tan matematizable y predecible. Lo mental parece poseer una serie de propiedades que, a primera vista, no son propias de lo físico:

b) **La conciencia.** La noción de "conciencia" no se refiere aquí a la mera conciencia (o autoconciencia) en contraposición a la "inconciencia", sino a los aspectos fenoménicos de las nuestras experiencias mentales. Esos aspectos fenoménicos se conocen generalmente como "**qualia**" (literalmente "cualidades", en latín), los aspectos cualitativos de cualquier estado mental, aquello en lo que consiste experimentar la vivencia de algo. Hay una diferencia entre describir un estado mental y tener ese estado mental.

a) **La subjetividad.** No conocemos nuestros estados mentales de la misma forma que un estado físico, no es necesario emplear nuestros sentidos para acceder a ellos. Mis estados mentales son internos, yo los experimento y cualquier otra persona no, solo accede a ellos de forma indirecta. A esto se le conoce como **autoridad de la primera persona**, la idea de que el conocimiento que cada uno de nosotros tiene de sus propios estados mentales es en cierto sentido privilegiado. La médico puede decirnos "no es posible que esto te duela", pero si nos duele, la médico no es quién para negarlo por mucho que sus conocimientos objetivos sobre aquello que nos está causando el dolor la lleve a concluir que es imposible que lo haga.

c) **Intencionalidad.** Nuestros estados mentales tratan acerca de algo, se representan algo, no se agotan en sí mismos (como ocurre con cualquier estado físico, que nunca es "acerca de"). Algunos estados mentales tienen de referir a algo distinto de sí mismos, de ser 'sobre' (como mis creencias) o 'de' (como mi miedo) o 'para' (como mis planes) algo. Estados mentales como creencias y deseos pueden expresarse como actitudes proposicionales porque se expresan como actitudes hacia ciertas proposiciones ("Creo que la filosofía es interesante" o "Deseo que la filosofía sea interesante" expresan mi actitud hacia la proposición "la filosofía es interesante").

d) **Racionalidad.** Si a la pregunta "¿Por qué dijo eso?", contesto con "Porque sus pulmones expelieron aire que al atravesar la laringe hizo vibrar las cuerdas vocales, y sus resonadores y articuladores modificaron dicho sonido", probablemente se me reproche que no he entendido la pregunta. La pregunta no indagaba en las causas del sonido emitido, sino en sus razones, pero la realidad física solo atiende a causas, no a razones, que son propias de la mente. Describir el comportamiento de las personas en términos de estados mentales intencionales es clasificar ese comportamiento como algo que debe explicarse apelando a razones. Las "razones" son un tipo de causa explicativa especial, pues se expresan en el lenguaje de la psicología y no de la física.

El contraste que existe entre estas propiedades propias de la mente y las típicas del mundo físico (los hechos físicos son objetivos y no subjetivos, se agotan en sí mismos en lugar de referir a algo distinto, y se explican mediante causas y no razones), es la expresión misma del problema mente-cerebro (alma-cuerpo en su versión clásica). Este problema tiene distintos frentes, pero todos tienen en común la distinción entre los fenómenos mentales y los fenómenos físicos, y una serie de premisas que impiden comprender cómo se relacionan ambos.

Resulta difícil entender cómo los pensamientos, los sentimientos, las percepciones, las acciones y otros fenómenos mentales se relacionan con impulsos eléctricos y reacciones químicas en el sistema nervioso humano. En nuestra vida cotidiana nos sentimos seres libres que actúan según sus creencias, deseos, esperanzas y temores. Cuando adoptamos el punto de vista científico, sin embargo, vemos el universo como un vasto mar de materia y energía, y a un nivel fundamental no hay diferencia entre los seres humanos y las rocas, los árboles o los animales, todos están hechos de partículas subatómicas que no tienen creencias ni deseos, esperanzas o temores. Nos enfrentamos así a dos imágenes de la vida humana: la imagen cotidiana y precientífica de nosotros mismos como seres libres y racionales con vidas mentales y morales, y la imagen científica de nosotros mismos como complejos sistemas bioquímicos. (Jaworski, 2011, p. 11-12, la traducción es mía)

Este contraste es especialmente evidente si, como quería Descartes, el alma y el cuerpo, las personas (como algo distinto de sus cuerpos) y sus organismos (como algo físico), son entes sustancialmente distintos. Si, por una parte, el “yo” es tan solo “algo que piensa” (cree, duda, sufre, ama...), que tiene estados mentales y no es por tanto físico sino intangible (y libre e inmortal, añade Descartes), y, por otra parte, el cuerpo tiene una superficie, ocupa un cierto espacio, tiene forma y movimiento e interactúa mediante contacto físico, entonces surgen varias preguntas: ¿Cómo puede lo intangible interactuar con lo tangible? ¿Cómo puede algo inextenso mover cosas extensas? ¿Cómo pueden el alma, el pensamiento, “golpear”, “empujar” o tener contacto alguno con el cuerpo? Este es básicamente el problema mente/cuerpo, cuando mente y cuerpo son tenidos por sustancias distintas, es difícil conciliar sus dos mundos, y de hecho Descartes no fue capaz de hacerlo (sugirió que la interacción tenía lugar en la glándula pineal del cerebro, pero el “dónde” no es el “cómo”). Podemos distinguir dos aspectos fundamentales en dicho problema:

a) **El problema de la causación mental.** Asumimos que nuestros estados mentales son responsables de nuestras acciones, esto es, que las acciones parecen ser sucesos físicos con causas mentales. No obstante, los sucesos físicos solo pueden ser desencadenados por otros sucesos físicos: si estimulamos directamente el sistema nervioso (mediante una sustancia química o una corriente eléctrica), podemos producir una respuesta física idéntica a la de nuestras acciones conscientes. Por lo tanto, dos causas distintas parecen coincidir en nuestras acciones: tienen una causa física, sucesos en nuestro sistema nervioso, y también tienen una causa mental, el deseo de realizarlas. ¿Cómo están relacionadas la causa mental de una acción y su causa física?

b) **El problema de la emergencia de lo psíquico a partir de lo físico.** La vida y la mente no han existido desde siempre, lo que sugiere que las condiciones físicas que precedieron su aparición tienen que ser de alguna forma responsables de que surgieran. Somos seres conscientes con estados mentales, sin embargo, estamos compuestos enteramente de partes no conscientes como moléculas. ¿Cómo surgen nuestras experiencias conscientes de estas interacciones físicas no conscientes? ¿Cómo logró la conciencia emerger en el curso de la historia del universo? ¿Y, ya puestos, cómo logra emerger a cada instante la conciencia en ti y en mí?

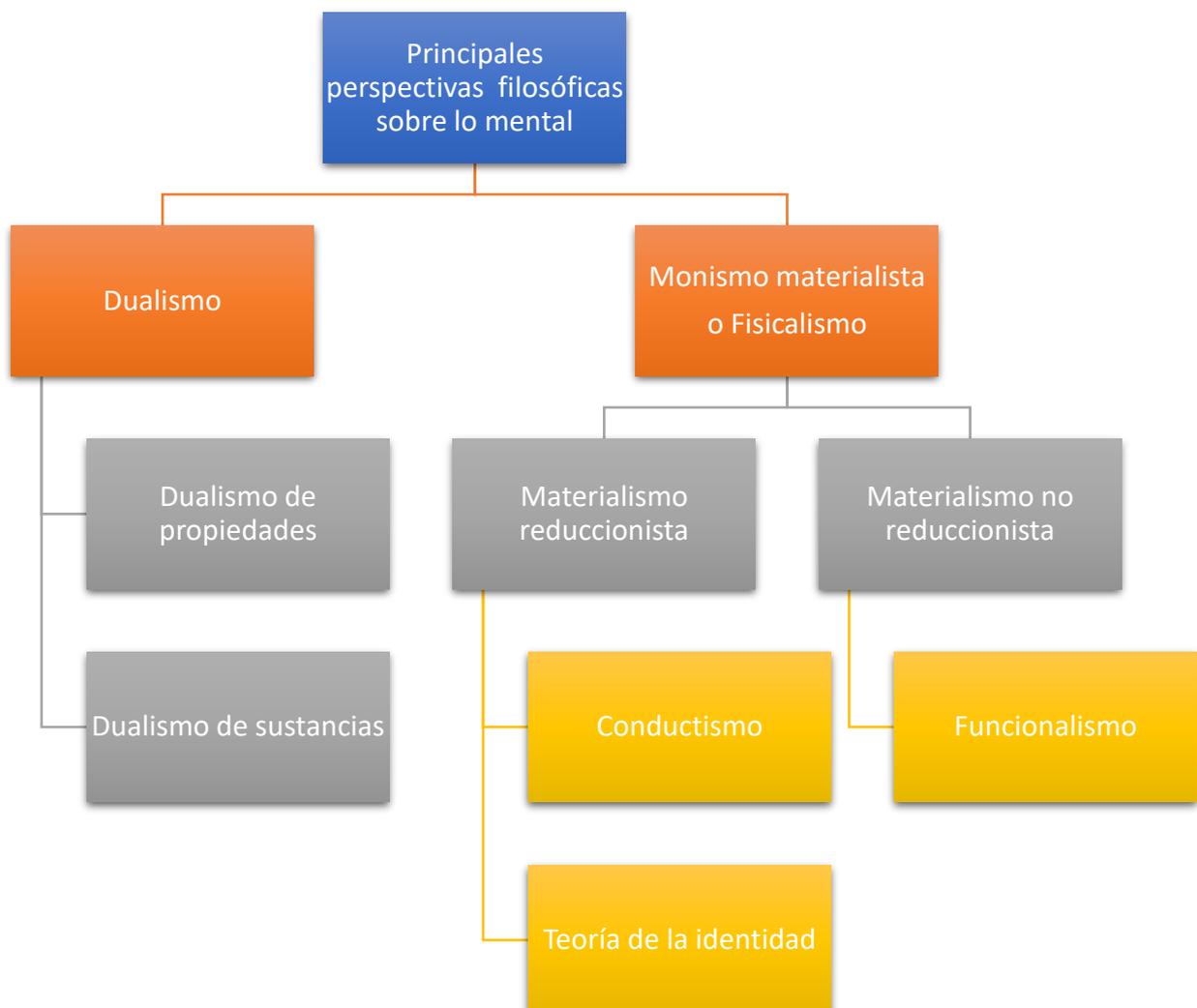
Estos problemas han sido abordados desde ópticas muy distintas. En el siguiente apartado veremos de forma resumida algunas de las distintas perspectivas que existen en filosofía de la mente, si bien nos centraremos en las teorías materialistas, esto es, aquellas que presuponen que no existe ninguna otra sustancia en la realidad aparte de la materia. El reto de estas teorías es dar cuenta de las propiedades de lo mental citadas más arriba, que parecen escapar al ámbito de lo físico (aunque por otra parte están libres del reto cartesiano de explicar la, por definición, imposible interacción

entre sustancias distintas). Dentro de la perspectiva materialista nos interesaremos de forma más exhaustiva por la solución a dicho reto propuesta por el funcionalismo: los estados mentales son estados funcionales.

II.1.2. Perspectivas ontológicas de lo mental

Aunque existen múltiples perspectivas sobre lo mental (como se puede ver en el diagrama posterior, que no obstante está lejos de ser exhaustivo), aquí vamos a centrarnos en las más influyentes en la filosofía de la mente de los últimos años, las materialistas, y en concreto de un tipo de materialismo no reduccionista, el funcionalismo. A grandes rasgos puede considerarse que las teorías fundamentales acerca de lo mental son las siguientes (corresponden al tercer nivel del diagrama):

- a) Alma (mente) y cuerpo (cerebro) son realidades sustanciales distintas, la persona es el alma (dualismo de sustancias).
- b) El alma (la mente) no existe como algo distinto del cuerpo (del cerebro), es física, la persona es el cerebro (materialismo reduccionista).
- c) El alma (la mente) no existe sin el cuerpo (el cerebro), pero los estados mentales no son estados físicos (dualismo de propiedades o materialismo no reduccionista).



El dualismo de sustancias (a) afirma que las personas y los cuerpos son distintos. Las personas son seres puramente mentales, no tenemos propiedades físicas. Los cuerpos, por otro lado, no tienen propiedades mentales porque son seres puramente físicos. No solo, entonces, hay dos tipos distintos de propiedades, mentales y físicas, sino que también hay dos tipos distintos de individuo o sustancias. Este dualismo es el de Platón y Descartes, aunque la palabra que usaban para "mente" o "persona" era "alma", y defendían que era inmortal, y por lo tanto diferente del cuerpo mortal. La reencarnación (en las religiones orientales y la filosofía de Platón) y la inmortalidad (en las religiones occidentales) son la principal motivación para defender el dualismo de sustancias: si somos seres no físicos unidos a un cuerpo podríamos ser capaces de una existencia independiente una vez el cuerpo está muerto. También, Descartes pensó que sólo una sustancia no física podría tener libre albedrío: si fuéramos parte del universo físico deberíamos estar completamente gobernados por las leyes físicas fundamentales, y por lo tanto no podríamos ser realmente libres. Además, los dualistas argumentan que es posible y concebible que se pueda existir sin un cuerpo, que entonces no sería una propiedad esencial de las personas (como sí lo serían el pensamiento o el libre albedrío) sino una propiedad accidental y, por lo tanto, ciertamente yo no sería mi cuerpo.

El monismo materialista o fisicalismo afirma que todo es físico, y todo puede ser descrito y explicado exhaustivamente por la física. Solo existe una sustancia en la realidad, la materia. Podemos considerar la filosofía de Aristóteles y de Hobbes ejemplos de materialismo (aunque muy alejados entre sí). Dentro del materialismo, no obstante, hay diferencias. El materialismo o fisicalismo reduccionista (b) defiende que el discurso psicológico ordinario (que emplea conceptos mentalistas tales como "creencias" o "deseos") corresponde de forma directa a categorías físicas. Es por tanto solo cuestión de tiempo que, según ampliamos nuestros conocimientos científicos, podamos reducir, esto es, traducir, los estados mentales a estados físicos. Entre las teorías reduccionistas destacan dos: el conductismo y la teoría de la identidad. Para el conductismo los estados mentales se reducen a conductas observables, las descripciones psicológicas son abreviaturas de descripciones físicas de comportamientos reales y/o potenciales. La teoría de la identidad defiende la reducción definitiva de los estados mentales a estados cerebrales. El principal motivo para defender el reduccionismo es el éxito científico de las descripciones físicas de hechos mentales. Históricamente las explicaciones no físicas de los acontecimientos naturales han fracasado (la epilepsia como "ser poseído por un demonio", por ejemplo) mientras que las descripciones físicas han tenido éxito (la epilepsia como un "trastorno neurológico que produce convulsiones epilépticas debido a la actividad excesiva y anormal de las células nerviosas corticales en el cerebro"). Es razonable pues apostar por una explicación física de los fenómenos mentales. Además, el materialismo no tiene el problema de la causación mental, ni de la emergencia de lo psíquico a partir de lo físico, pues no hay una segunda sustancia que deba interactuar con lo físico.

Al margen de estas posturas dualistas o materialistas estrictas, existen filósofos que defienden un dualismo de propiedades pero no de sustancias, o un materialismo no reduccionista (c). Por una parte, los partidarios del dualismo de propiedades defienden que las propiedades mentales y físicas son distintas (como afirma el dualismo de sustancias) y que el mismo individuo puede tener propiedades de ambos tipos (como afirman la mayoría de versiones del materialismo). Las formas más populares de dualismo de propiedades son el emergentismo (las propiedades mentales emergen de la organización de componentes con propiedades físicas y no son meras descripciones abstractas de procesos de un nivel inferior, sino una clase distinta de propiedades) y el epifenomenalismo (las propiedades mentales emergentes son meramente subproductos causales de ciertos procesos físicos y son causalmente inertes). Ambos afirman pues que algunos organismos tienen propiedades mentales y físicas y que las propiedades mentales son generadas o producidas por interacciones físicas en esos organismos físicos. Por otra parte, los materialistas no reduccionistas consideran que el hecho de que los acontecimientos mentales correspondan a

los sucesos físicos no implica necesariamente que el lenguaje de lo mental deba reducirse al lenguaje de lo físico. Sostienen que la física satisface algunas de nuestras necesidades descriptivas y explicativas, pero no todas, y el vocabulario de lo mental es útil desde el punto de vista descriptivo y explicativo. Dos teorías populares de materialismo no reduccionista son el funcionalismo y el monismo anómalo. Aquí vamos a ocuparnos en detalle de la primera, partiendo de su crítica a las grandes teorías por entonces dominantes en la filosofía de la mente, los ya mencionados conductismo y teoría de la identidad.

II.1.3. El funcionalismo

En su seminal artículo de 1967 "La naturaleza de los estados mentales", Hilary Putnam plantea la posibilidad de que el fisicalismo no implique necesariamente reduccionismo, esto es, que sea posible conservar el lenguaje propio de lo mental sin por ello suponer propiedades no físicas o, en cualquier caso, entidades no físicas. Putnam trata pues de ofrecer una solución teórica compatible al mismo tiempo con nuestros conocimientos científicos acerca del cerebro y el comportamiento, y con conceptos psicológicos tales como "creencia" o "deseo". Dicha solución recibe el nombre de "funcionalismo" por el hecho de que los estados mentales son entendidos como estados funcionales.

En el citado artículo, Hilary Putnam se centra en un estado mental que parecemos comprender relativamente bien, el dolor, y se pregunta: "¿Es el dolor un estado cerebral?" (o bien: "La propiedad de tener un dolor en el tiempo t , ¿es un estado cerebral?"). (Putnam, 1967, p. 1) La respuesta de Putnam es que "no", y del mismo modo responde a la cuestión de si sería una disposición de conducta. Así, Putnam critica las dos teorías entonces dominantes dentro del materialismo antropológico: la teoría de la identidad, según la cual los estados mentales son estados cerebrales, y el conductismo, según la cual los estados mentales son disposiciones conductuales. Dichas teorías defendían, de facto, la sustitución de nuestro vocabulario mentalista por el lenguaje técnico de la neurología o de la psicología conductista, respectivamente, y lo que Putnam está defendiendo es que tal reducción es imposible e innecesaria, que el materialismo o fisicalismo no implica necesariamente reduccionismo. Esta es, resumida, la propuesta de Putnam (primero se ofrecen las definiciones dadas por Putnam para poder expresar con propiedad su hipótesis, y a continuación el desarrollo de la hipótesis propiamente dicha):

Propongo la hipótesis de que el dolor, o el estado de tener dolor, es un estado funcional de todo un organismo. [...]

- *[U]n **Autómata Probabilista** se define de manera similar a la de una Máquina de Turing, excepto en que se permite que las transiciones entre "estados" se den con varias probabilidades en lugar de ser "deterministas", [...] permite "entradas sensoriales" y "salidas motoras", es decir, la Tabla de la Máquina específica, para cada combinación posible de un "estado" con un conjunto completo de "entradas sensoriales", una "instrucción" que determina la probabilidad del "estado" siguiente, así como la probabilidad de las "salidas motoras". [...]*

- *Puesto que un sistema empíricamente dado puede ser a la vez una "realización física" de muchos Autómatas Probabilistas diferentes, introduzco la noción de **Descripción de un sistema**. Una descripción de S , en donde S es un sistema, es cualquier enunciado verdadero que dice que S posee los estados diferentes S_1, S_2, \dots, S_n , los cuales están relacionados entre sí, y con las salidas motoras y las entradas sensoriales, mediante las probabilidades de transición dadas en tal y cual Tabla de Máquina.*

- *La Tabla de Máquina mencionada en la descripción se llamará la **Organización Funcional** de S relativa a esa Descripción, y el S , tal que S se halla en el estado S_i en un momento dado, se llamará el Estado Total de S (en el momento) relativo a esa Descripción. [...]*

La hipótesis de que "ser un dolor es ser un estado funcional de un organismo" puede ahora explicitarse de manera más exacta de la siguiente forma:

(1) Todos los organismos capaces de sentir dolor son Autómatas Probabilistas.

(2) Todo organismo capaz de sentir dolor posee por lo menos una Descripción de un cierto tipo (esto es, ser capaz de sentir dolor es poseer una clase especial de Organización Funcional).

(3) Ningún organismo capaz de sentir dolor puede descomponerse en partes que posean separadamente Descripciones del tipo de las mencionadas en (2).

(4) Para cada Descripción de la clase mencionada en (2), existe un subconjunto de entradas sensoriales tal que un organismo con dicha Descripción siente dolor cuando, y sólo cuando, algunas de sus entradas sensoriales están en ese subconjunto.

[...] [L]a Organización Funcional debe incluir algo semejante a una "función de preferencia", o por lo menos a un ordenamiento parcial de preferencia, y algo semejante a una "lógica inductiva" (esto es, la Máquina tiene que ser capaz de "aprender de la experiencia"). Parece necesario, además, exigir que las máquinas posean "sensores de dolor", esto es, órganos sensoriales que normalmente indiquen daños al cuerpo de la Máquina, o temperaturas, presiones, etcétera, que representen un peligro, y que dichos sensores transmitan un subconjunto especial de las entradas, el subconjunto al que nos referimos en (4). (Putnam, 1967, págs. 4-5)

Como vemos, el funcionalismo putnamiano parte del concepto de Máquina de Turing², considerando que las descripciones psicológicas son como las descripciones de una máquina de Turing: descripciones abstractas que postulan entradas, salidas y estados internos que correlacionan ambos (en particular creencias y deseos). En su versión primigenia, como "funcionalismo de máquina", nuestro cerebro sería un mero hardware y nuestra mente un software capaz de correlacionar las entradas a un sistema con sus salidas de manera apropiada (satisfaciendo la prueba de Turing), lo que ha llevado a algunos autores a considerar el funcionalismo como sinónimo de lo que Searle (véase Searle 2001) llama "IA fuerte", algo que no es necesariamente cierto, especialmente de otras variedades del funcionalismo como el psicofuncionalismo de Jerry Fodor, el funcionalismo analítico de David Lewis o el homuncular de Daniel Dennett y William Lycan.

Los conceptos clave del funcionalismo son el de propiedad funcional, el de superveniencia (que es posterior al artículo de Hilary Putnam, pero está implícito en su teoría) y el de realización:

Una propiedad funcional F de un objeto es una propiedad que ese objeto posee exclusivamente en virtud de su aptitud para cumplir cierto papel causal en un determinado contexto. De todos aquellos objetos que tengan esta aptitud puede decirse que tienen la propiedad funcional F y pueden por ello clasificarse como Fs, aun cuando difieran ampliamente en sus propiedades físicas de primer orden. (Moya, 2006, p. 99)

Por sus propiedades funcionales un e-book, un libro en braille, un papiro, un códice o un audio-libro son igualmente libros, aunque difieran en su soporte físico. Todos ellos existen físicamente, tienen propiedades funcionales porque tienen propiedades físicas, pero aquellas no son idénticas a estas.

Dos objetos pueden ser físicamente distintos y funcionalmente equivalentes. Lo que no puede suceder, sin embargo, es que dos objetos sean físicamente idénticos y difieran en sus propiedades funcionales [...]. En este sentido, las propiedades físicas son más básicas que las propiedades funcionales. Podemos denominar esta relación de dependencia metafísica entre ambos grupos de propiedades una relación de superveniencia. (Moya 2006, p. 100)

² Véase Turing (1950).

La superveniencia³ es la alternativa funcionalista a la identidad. Los estados mentales no son idénticos a los estados cerebrales, sino que supervienen a ellos, esto es, si difieren en algún aspecto mental, entonces también en lo físico, y no es posible que sean físicamente idénticos y distintos mentalmente. Lo que sí es posible es que sean mentalmente idénticos y físicamente distintos, esta es la base del funcionalismo, que un mismo estado mental puede *realizarse* en distintos estados físicos, de la misma forma que “En un lugar de la Mancha, de cuyo nombre no quiero acordarme...” puede realizarse en tinta electrónica, en tinta impresa sobre papel, o como ondas sonoras emitidas por el altavoz de un dispositivo electrónico. Eso sí, sin soporte físico (sin bites, ni papel, ni ondas) no hay libro, no hay principio del *Quijote*, de ahí que el funcionalismo sea una forma de materialismo, pero no reduccionista, la realizabilidad múltiple de los estados funcionales no permite identificarlos (y reducirlos, por tanto) con estados cerebrales o con disposiciones conductuales. De hecho, la realizabilidad múltiple de los estados mentales es la mejor razón para entenderlos como estados funcionales antes que como estados cerebrales o disposiciones conductuales. Así, el auténtico fundamento del funcionalismo de Putnam es el argumento de la realizabilidad múltiple, que es el que le permite responder en su artículo de 1967 que el dolor no es un estado cerebral.

Consideremos lo que el teórico de los estados cerebrales tiene que hacer para dar validez a sus afirmaciones. Tiene que especificar un estado fisicoquímico tal que cualquier organismo (no solamente un mamífero) siente dolor si y sólo si (a) posee un cerebro con una estructura fisicoquímica adecuada y (b) su cerebro está en ese estado fisicoquímico. Esto significa que el estado fisicoquímico en cuestión tiene que ser un estado posible de un cerebro de mamífero, de un cerebro de reptil, de un cerebro de molusco (los pulpos son moluscos y ciertamente sienten dolor), etcétera. Al mismo tiempo, tiene que ser un estado que no sea posible (físicamente posible) para el cerebro de ninguna creatura físicamente posible que no pueda sentir dolor. [...]

No es completamente imposible que se encuentre un estado semejante. [...] Sin embargo, esta es ciertamente una hipótesis ambiciosa.

Finalmente, la hipótesis se torna aún más ambiciosa cuando caemos en la cuenta de que el teórico de los estados cerebrales no solamente dice que el dolor es un estado cerebral; lo que le interesa mantener, por supuesto, es que todo estado psicológico es un estado cerebral. Entonces, si podemos encontrar aunque sea un solo predicado psicológico que pueda aplicarse claramente tanto al mamífero como al pulpo (digamos, "hambriento"), pero cuyo "correlato" fisicoquímico sea diferente en los dos casos, la teoría del estado cerebral se derrumba. Me parece sumamente probable que podamos hacer eso. (Putnam 1967, p. 6)

Esto es, ciertamente dos seres con estados cerebrales idénticos coincidirán en su estado mental, pero un mismo estado mental puede realizarse en estados cerebrales distintos. No solo es sumamente probable que distintas especies puedan compartir estados mentales sin compartir estados cerebrales (dada la diferente naturaleza de sus cerebros), sino que es también probable que no existan entre personas distintas, estados cerebrales estrictamente idénticos a todos los niveles, y defender la identidad entre estados cerebrales y estados mentales implicaría que existen tantos estados mentales como instancias concretas de dichos estados en cada cerebro humano. No existiría el estado mental “tristeza”, sino tristeza en el sujeto S_1 , tristeza en el sujeto S_2 , en el $S_3... S_n$. Es más, habría probablemente que considerar que un mismo sujeto nunca está en el mismo estado mental, aun cuando se trate del mismo estado mental a todos los efectos salvo el del estado cerebral, habría que hablar de tristeza en el tiempo $T_1, T_2, T_3... T_n$. La hipótesis más sensata es considerar que un mismo estado mental es realizable en múltiples estados físicos.

³ Este concepto se debe a Donald Davidson, que lo introduce en su artículo de 1980 “Mental events”.

¿Y qué hay de la idea conductista de que un estado mental es una disposición conductual? Suena suficientemente semejante a un estado funcional. No obstante, la caracterización de los estados mentales en términos conductuales es, por una parte, vaga y, peor aún, circular, cosa que no ocurre con su caracterización en términos funcionales. “La dificultad (de hecho, parece ser más que una simple ‘dificultad’) de especificar la disposición conductual requerida excepto como ‘la disposición de X a comportarse como si X sintiera dolor’, es desde luego la más importante.” (Putnam 1967, p. 7) El funcionalismo admite que no es necesario traducir el estado mental en una conducta posible o esperable, el estado mental tiene entidad por sí mismo, tiene sentido en el contexto del “Estado total (es decir, de otros valores, creencias, etcétera)” del organismo, y no implica necesariamente (o no tiene sentido por) su potencial conversión en conducta.

La teoría de la disposición conductual no sólo parece irremediablemente vaga, sino que parece claramente falsa si la "conducta" a la que se refiere es la conducta periférica y los estímulos relevantes son los estímulos periféricos (v. gr. no se dice nada acerca de lo que el organismo haría si se le opera el cerebro). Por ejemplo, dos animales con todos los nervios motores cortados tendrán la misma conducta presente y potencial (esto es, ninguna, por así decirlo); pero si uno de ellos tiene cortadas las fibras del dolor y el otro las tiene intactas, entonces uno de ellos sentirá dolor y el otro no. (Putnam 1967, p. 8)

Podemos encontrar un contraejemplo semejante, esto es, contrario a la identidad entre estados mentales y disposiciones conductuales en la ficción *La naranja mecánica*, que cuenta como un experimento mental en que vemos una clara disociación entre deseos y creencias y su traducción en conductas o la potencialidad de llevar a cabo dichas conductas. La terapia de aversión “Ludóvico” ciertamente condiciona conductualmente al protagonista de *La naranja mecánica*, Álex, pero ello no afecta (o no completamente) a sus estados mentales. Aunque tras la terapia Álex sea incapaz de disfrutar con la Novena Sinfonía de Beethoven, no por ello sus estados intencionales “me gusta la Novena Sinfonía de Beethoven”, “la Novena Sinfonía de Beethoven es mi composición favorita” han pasado a ser falsos; ni ha aprendido que la violencia es mala, aunque sea físicamente incapaz de cometer acto violento alguno. La repulsión generada por el Himno de la Alegría o por la percepción de la violencia no es consciente. Dado que es lógicamente posible poseer un determinado estado mental, pero que este no corresponda con la disposición conductual asociada, entonces puede concluirse que no cabe reducir el primero a la segunda.

En síntesis, en su artículo Putnam, por una parte, ofrece un sólido argumento contra el materialismo reduccionista, el de la realizabilidad múltiple, que puede resumirse del siguiente modo:



Y, por otra parte, Putnam propone una teoría concreta no reduccionista, el funcionalismo. El filósofo ofrece una hipótesis plausible de cómo podríamos entender los estados mentales: las propiedades mentales son propiedades funcionales de determinados seres, y cabe explicar tanto el carácter intencional como consciente de los estados mentales en dichos términos. No todos los filósofos de la mente materialistas han admitido el argumento de la realizabilidad múltiple, y muchos menos han abrazado el funcionalismo, no obstante, a raíz de las ideas de Putnam (y otros filósofos como Donald Davidson), el materialismo no reduccionista se ha convertido en la perspectiva principal dentro del materialismo.⁴

Profundizando algo más en la teoría funcionalista, hemos de entender que lo que hace de un estado mental que sea el que es, es estar relacionado causalmente de forma adecuada con otros procesos o estados mentales del sujeto, lo que implica que sea susceptible de cambio de acuerdo con estímulos sensoriales, respuestas motoras, así como por su interacción con otros estados mentales. “Estar en un estado mental equivale a tener un conjunto de estados internos de primer orden que están relacionados entre sí de manera que satisfagan una descripción funcional” (Jaworski 2011, p. 141), ocupen un determinado papel causal. El lenguaje de lo mental, el relato psicológico, no es como el de las teorías científicas naturalistas, sino abstracto, recurre a idealizaciones abstractas (primer orden) de propiedades físicas concretas (segundo orden), cuyas concreciones son ignoradas, y por ello es realizable en distintos estados físicos.

Tener dolor, por ejemplo, puede definirse como estar en un estado concreto de primer orden correlacionado con pinchazos, quemaduras y abrasiones, con contracciones de dolor y quejidos, y estados internos tales como irritación y disgusto. Decir que el cerebro de Alexander está realizando en este momento un estado de dolor, por tanto, equivale a decir que su cerebro está en un estado concreto de primer orden que está relacionado con otros estados de primer orden de manera que satisface la condición asociada al dolor. (Jaworski 2011, p. 141)

Aunque esta descripción pueda parecer en principio idéntica al conductismo, es necesario subrayar ese “estados internos” en que insiste el funcionalismo. Así lo hace Ned Block en “Las dificultades del funcionalismo”, donde hace una caracterización general del funcionalismo distinguiéndolo específicamente del conductismo:

[C]ada tipo [type] de estado mental es un estado que consiste en una disposición a actuar de ciertas maneras y a tener ciertos estados mentales. [...] El conductismo identifica los estados mentales con disposiciones a actuar de ciertas maneras en ciertas situaciones de input. Pero [...] desear G como meta, no puede identificarse con, digamos, la disposición a hacer A en circunstancias de input en las cuales A conduce a G, puesto que, después de todo, el agente podría no saber que A conduce a G y de este modo podría no estar dispuesto a hacer A. El funcionalismo reemplaza a los “inputs sensoriales” conductistas por “inputs sensoriales y estados mentales”, y el funcionalismo reemplaza las “disposiciones a actuar” conductistas por “disposiciones a actuar y tener ciertos estados mentales”. (Rabosi 1995, pp. 105-106)

La mejor herramienta para entender muchos de los conceptos del funcionalismo es el ensayo de Daniel Dennett en forma de relato de ciencia ficción “¿Dónde estoy yo?”, que además permite, de forma independiente al estudio del funcionalismo, plantear interesantes cuestiones en filosofía de

⁴ Nótese, no obstante, que entender los estados mentales como estados funcionales no implica una ontología materialista. De hecho, la mayor parte de funcionalistas son materialistas (porque hace innecesario el coste extra de asumir la existencia de una sustancia no física, fuera del relato científico), pero el funcionalismo es compatible con el dualismo. Es, desde el punto de vista ontológico, neutral.

la mente.⁵ En dicho relato el propio Dennett es el protagonista y nos cuenta en primera persona las aventuras y desventuras de su yo tras una serie de trasplantes de cerebro (o de cuerpo, según se mire). El filósofo nos relata cómo fue reclutado para una peligrosa misión ultrasecreta, y que para realizarla, por mayor seguridad, en primer lugar construyen un relato exhaustivo de su mente (creencias, deseos, valores, principios, gustos, temores...), y en segundo lugar extraen su cerebro de su cuerpo, que no obstante sigue siendo controlado por aquel a distancia mediante unos transmisores punteros que sustituyen todas las conexiones nerviosas mente-cuerpo. El caso es que en el decurso de su misión el cuerpo de Dennett fallece, y el cerebro de Dennett queda sumido en la oscuridad de sus pensamientos hasta que le conectan un nuevo cuerpo. Descubre entonces, también, que, perfectamente sincronizado con su cerebro orgánico, funciona un cerebro computerizado que han fabricado programándolo según el relato exhaustivo que Dennett facilitó de su mente. A su vez, Dennett posee un interruptor que le permite activar cualesquiera de sus dos cerebros, que están conectados de forma idéntica a su cuerpo y reciben por tanto exactamente los mismos estímulos y viven las mismas experiencias, aunque solo aquel que está activo puede realmente mover el cuerpo (si bien, ambos cerebros, al funcionar en paralelo y estar en estado funcionales idénticos, recibiendo entradas idénticas, siempre producirían de forma indistinguible la misma respuesta). Ambos cerebros, el orgánico y el digital, viven cómo si fuesen dueños del cuerpo hasta que, sin que se sepa cómo, se da un pequeño retraso en la recepción de estímulos externos en el caso de uno de ellos y, por efecto dominó, este acaba sintiendo que ha perdido control sobre el cuerpo porque esa mínima diferencia de partida termina generando estados mentales distintos de los de su contraparte, lo que sugiere que, al final del relato ya no hay una única persona, sino dos.

En este relato aparecen muchos de los tropos del funcionalismo: vemos que no hay identidad del yo con el cerebro (si solo el cerebro de Dennett fuera encarcelado por cometer un delito, pero no el cuerpo, sin duda podría decirse que Dennett es una persona libre); que la interacción con el entorno, el mecanismo de entradas y salidas de estímulos y de acciones motoras, el poseer un cuerpo, el deambular y vivir experiencias, la forma en que se producen los cambios físicos, sería relevante para poseer una mente; vemos que Yorick puede ser intercambiado por Hubert; vemos que, mientras no se altere la interacción causal con el entorno, no hay cambios en la mente, pero que una mínima diferencia en dicha interacción produciría un cambio que nos llevaría a hablar de una mente nueva. Y todo ello desde un punto de vista estrictamente materialista y, no obstante, sin reduccionismo (o pruébese a reducir el concepto de “punto de vista”).

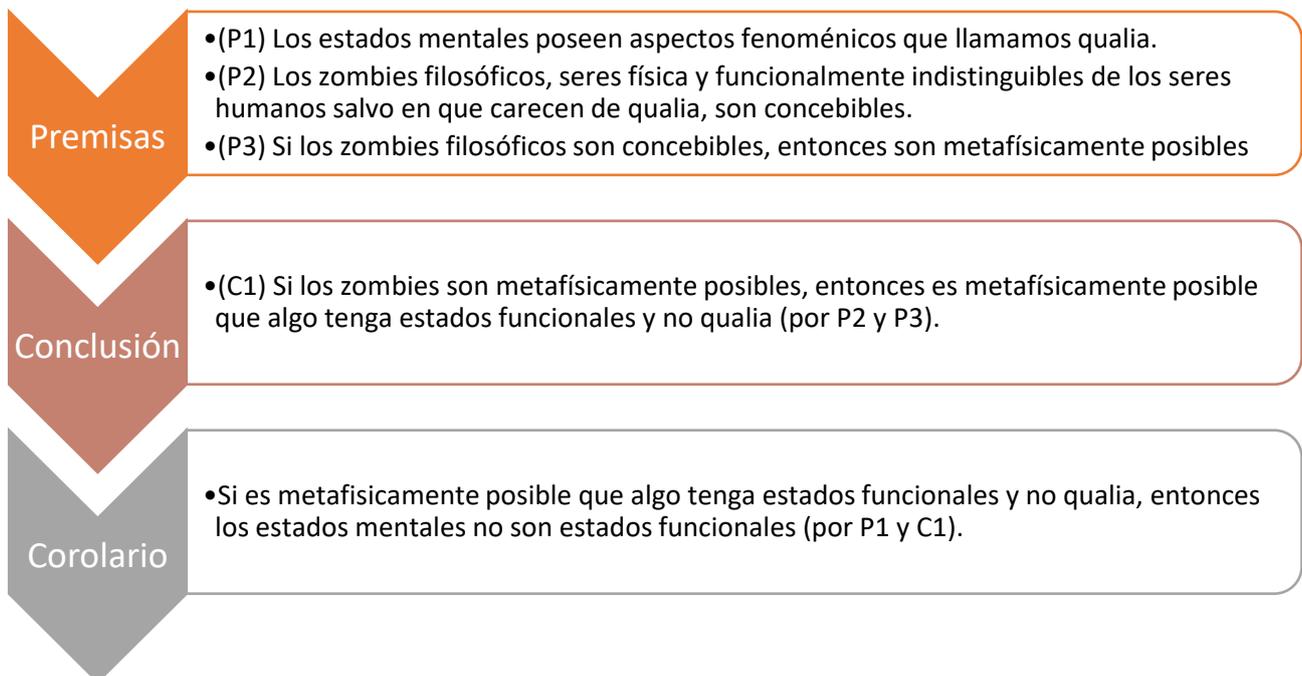
II.1.4. Algunas críticas al funcionalismo: los qualia y la habitación china

A pesar de (o tal vez, precisamente, a causa de) su éxito, el funcionalismo se ha enfrentado a numerosas críticas. Vamos a resumir aquí los principales argumentos en su contra. En primer lugar, el funcionalismo es sensible a las críticas habituales que parten del concepto de “qualia” a las que se enfrenta el materialismo; además, se le reprocha al funcionalismo su excesivo liberalismo, pues podría llevar a atribuir estados mentales a sistemas que resulta más o menos obvio que no los tienen; y, por fin, el funcionalismo ignoraría el carácter semántico de la comprensión, de los estados mentales, postulando mentes puramente sintácticas. Hay muchas versiones de estas críticas, aquí vamos a considerar un argumento dentro de cada una de estas estrategias: el de los zombis filosóficos de David Chalmers, el de la habitación china de Ned Block y el de la habitación china de

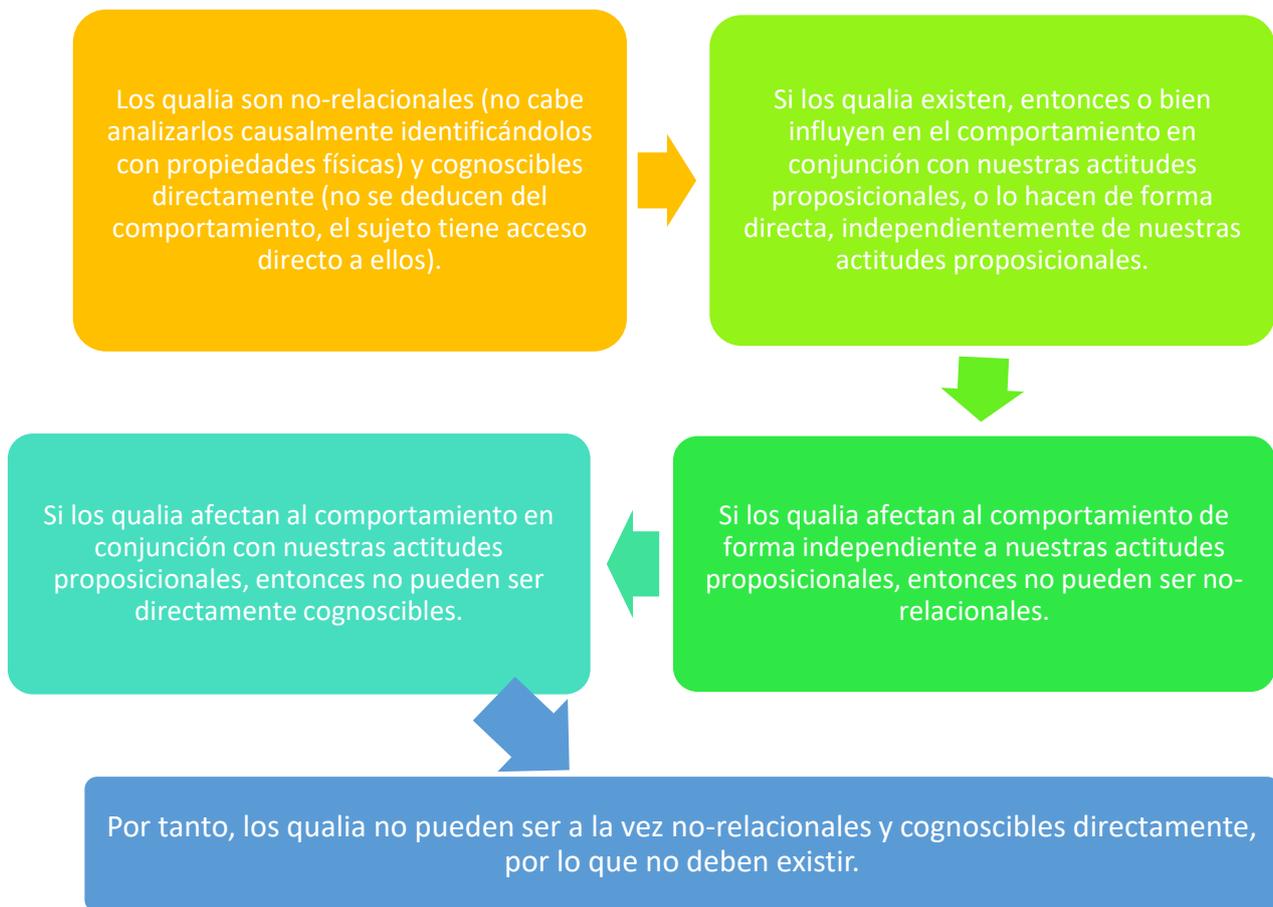
⁵ La actividad prevista para los alumnos respecto a este tema consiste precisamente en plantearles dichas cuestiones al hilo de la lectura del texto de Dennett, que puede encontrarse en este enlace: <https://documents.es/document/donde-estoy-yo-dennett.html>

John Searle. Por fin, mencionaremos brevemente algunos contraargumentos ofrecidos desde posiciones funcionalistas (en concreto por Daniel Dennett).

Zombies filosóficos. Este argumento no es específico contra el funcionalismo, pero en su libro *La mente consciente: en busca de una Teoría Fundamental* de 1996 David Chalmers desarrolla su experimento mental de los zombies filosóficos, un argumento que pertenece a la familia de argumentos anti-fisicalistas basados en el carácter subjetivo de la mente y presencia o ausencia de qualia (el ejemplo más destacado de esta categoría de argumentos sería el de María y la habitación en blanco y negro de Frank Jackson). Según David Chalmers un zombie filosófico es un cuerpo humano idéntico a un humano en todos los aspectos posibles excepto en uno, la presencia de una experiencia fenoménica consciente. Es concebible la idea de estar rodeado por múltiples seres semejantes a mí, que actúan exactamente como seres humanos, esto es, que actúan exactamente como si tuvieran estados conscientes, y que no obstante carecen de dichos estados mentales que yo sí tengo. Y la idea de Chalmers es que si es concebible, entonces es un estado de cosas posible, esto es, podría haber zombies filosóficos, y si los hubiera, serían indistinguibles en términos funcionalistas de los humanos, por lo que, en términos funcionalistas, habría que atribuirles una mente, lo cual es incoherente, pues partimos del hecho de que la subjetividad y los qualia son fundamentales para definir la mente. El argumento puede descomponerse del siguiente modo:



Hay varias formas en que el funcionalismo puede enfrentarse a este argumento. Por una parte, criticando la premisa (P2) defendiendo que el que algo sea concebible no implica que sea posible. Por otra, defendiendo que los zombies filosóficos no son consistentemente concebibles (tal vez los zombies filosóficos nos parecen concebibles debido a nuestra ignorancia, y si supiéramos definir lo que llamamos qualia en términos físicos, entonces entenderíamos que el zombie filosófico es un concepto contradictorio). Por fin, negando la existencia de qualia, o al menos su existencia en los términos en que suele definirse, o si no en aquellos términos en que parece oponerse al funcionalismo. Daniel Dennett en *La conciencia explicada* ha defendido las dos últimas estrategias, su argumento contra los qualia es como sigue:



La nación china. Aunque existen diversas versiones del mismo argumento, e incluso su versión original puede verse en el relato de ciencia ficción *El juego* de Anatoly Dneprov, aquí lo estudiaremos en la forma en que lo formula Ned Block en “Las dificultades del funcionalismo”. Un corolario fundamental del argumento de la realizabilidad múltiple es que no solo los seres humanos o los seres con cerebros semejantes a los humanos podrían tener estados mentales, sino cualquier sistema con estados funcionales (por ejemplo, a alienígenas con una configuración física muy distinta de la nuestra). Si bien esto permite discutir con cierto sentido si una máquina, en sentido estricto, podría pensar, el problema es que una máquina de Turing no determinista podría ser muchas cosas, y a muchas de ellas nos resultaría absolutamente contraintuitivo atribuirle estados mentales, por mucho que tenga estados funcionales. En ocasiones se llama a este argumento el argumento del “liberalismo” porque, si bien la teoría de la identidad parece demasiado conservadora atribuyendo estados mentales, el funcionalismo es demasiado liberal concediéndolos:

Sea T una descripción funcional de las relaciones entre estados E_1, E_2, \dots, E_n – una descripción realizada por los estados del cerebro de Alexander. De acuerdo con T , por ejemplo, el estado E_1 conlleva el estado E_2 con una probabilidad del 0,73 dado el impulso I_{15} ; y los estados N_5, N_{67} , y N_4 en el cerebro de Alexander proporcionan un modelo de dichas relaciones: el estado cerebral de Alejandro N_5 conlleva el estado N_{67} con una probabilidad del 0,73 dado el impulso neuronal B_4 . [...] Imaginemos, no obstante, que se le ha diagnosticado a Alexander una enfermedad neurológica degenerativa [...] [y] los médicos reemplazan sus neuronas con prótesis de silicona. Cada prótesis de silicona funciona exactamente de la forma en que lo hace una neurona de Alexander [...]. Tras la cirugía, no son ya los estados cerebrales de Alexander los que realizan T , sino los estados de sus circuitos de silicona. Si el funcionalismo es cierto, entonces dicha situación es posible [...]. Imaginemos ahora otro posible escenario. El gobierno de China, muy impresionado por los resultados del tratamiento

de Alexander, lo abordan con una nueva propuesta. [...] [L]os funcionarios del gobierno quieren reemplazar las prótesis de silicona con algo distinto: personas, los ciudadanos de China. Cada persona físicamente capaz de China estará provista con un radio-transmisor y un radio-receptor que le permitirá interactuar con otras personas en China exactamente de la forma en que lo hacen las prótesis de silicona de Alexander en ese momento. Justo de la misma forma en que el sistema de circuitos de silicona de Alexander pudo realizar la organización funcional originariamente realizada por su cerebro, así también los ciudadanos de China realizarán la organización funcional ahora realizada por el sistema de circuitos de silicona de Alexander. [...]



China Brain por Danko Georgiev en la English Wikipedia
CC BY-SA 3.0, <https://commons.wikimedia.org/w/index.php?curid=17244363>

Si el funcionalismo es cierto, y el nuevo sistema consistente en los ciudadanos de China funciona exactamente de la misma forma que el sistema de circuitos de silicona y el cerebro de Alexander, entonces Alexander no habrá sufrido cambio mental alguno, pues, si el funcionalismo es cierto, la materia no importa. [...] [N]o obstante, esto es absurdo. Es absurdo suponer que los estados mentales de Alexander podrían ser realizados por un sistema de este tipo. (Jaworski 2011, pp. 150-151, he optado por traducir “input” como “impulso”)

Podemos sintetizar este argumento de la forma siguiente:



Este *modus tollens* es obviamente válido, luego el funcionalismo solo puede defenderse de este argumento negando alguna de las premisas.

En primer lugar, los funcionalistas pueden considerar que en la primera premisa el enunciado “sistemas como la nación china pueden tener estados mentales” no se sigue de la verdad del funcionalismo. Por una parte, porque el funcionalismo requeriría la recepción de un tipo concreto de impulsos que serían “sensoriales” de una forma restringida, aplicándose solo al tipo de estímulos que estamos dispuestos a considerar sensibles en cuanto a que implican reacciones en el orden del agrado y desagrado, placer y dolor (pero entonces el funcionalismo corre el riesgo de convertirse en un materialismo reduccionista de corte conductista). Por otra parte, los funcionalistas pueden defender que los estados mentales se realizan solo en sistemas “teleológicamente” organizados, esto es, que sirven a un propósito común, no son ciegos, implican deliberación y elección. La nación china no sería un sistema teleológico.

En segundo lugar, en una suerte de huida hacia delante, los funcionalistas pueden considerar que, de hecho, la nación china, en estas circunstancias, sería una mente, por muy contraintuitivo que nos resulte, es propio de una teoría revolucionaria que el nuevo paradigma resulte inconmensurable.⁶ Es más, algunos funcionalistas podrían decir que si nos tomáramos el experimento mental en serio, esto es, hiciésemos el esfuerzo de construir un relato que fuera más allá de la mera exposición de la situación, entonces nos daríamos cuenta de que hablar de estados mentales de la nación china no es ni siquiera tan contraintuitivo (o no más que hablar de estados mentales en general).

Habitación china. Sin duda el argumento más famoso en contra del funcionalismo es el de John Searle de la habitación china.



Searle's Chinese room 1 por Alexander Michels. Disponible en: https://www.researchgate.net/figure/Searles-Chinese-Room-1_fig1_334363118

En dicho experimento mental, Searle nos pide, para preguntarnos si una máquina tiene estados mentales, que nos imaginemos siendo la máquina. Supongamos, en primer lugar, que no hablamos ni entendemos chino, y que nos hallamos en una habitación con dos cajones con múltiples símbolos chinos (todos aquellos de los que consta el idioma chino), un libro de reglas en inglés que nos explica cómo asociar símbolos chinos (sin traducirnoslos) entre sí a partir de otros símbolos chinos (todas las asociaciones posibles), y por fin, un par de ranuras. Por una de las ranuras, que consideraremos de entrada, nos llegan papeles con símbolos chinos que han de ser tenidos como

⁶ Para más información a este respecto puede verse la unidad “Micro o macro, depende de la lente”.

un estímulo por el cual nosotros debamos, aplicando las reglas del libro de instrucciones, asociar los símbolos chinos pertinentes de los cajones y plasmarlo en un papel que, a su vez, sacaremos de la habitación por la segunda ranura que consideraremos de salida. Visto desde fuera, esos papeles que nos introducen hablantes competentes de chino son preguntas en chino, y los papeles que nosotros extraemos de la habitación, son respuestas consistentes con las preguntas y expresadas en un correcto chino, gracias a las instrucciones del manual que hay dentro de la habitación (nuestro programa). Según Searle, desde el punto de vista externo, aparentemente yo estoy hablando chino, pues estoy siendo capaz de responder competentemente a las preguntas que se me formulan, pero, *ex hypothesi*, yo no hablo ni entiendo el chino. Por así decir, yo supero la prueba de Turing de chino, pero es obvio que no comprendo lo que hago, funcionalmente lo hago, pero en sentido estricto no es así. La idea es que las máquinas solo podrían pensar de la forma en que un hablante estaría hablando chino en la habitación china, solo a nivel sintáctico, no semántico, las reglas de manipulación de símbolos son puramente formales. Una máquina podría superar el test de Turing sin que esto implicara que piensa, que tiene una mente como la humana, basta con ser capaz de asociar símbolos competentemente, pero, nos dice Searle, esto no basta para tener estados mentales. En las conclusiones a su artículo “Mentes y cerebros sin programas” (en Rabosi 1995, pp. 442-443), Searle sintetiza las premisas de las que parte su experimento mental y las conclusiones que se siguen de las mismas, quedando su argumento de la forma siguiente:

Premisas

- (P1) Los cerebros causan mentes.
- (P2) La sintaxis no es suficiente para la semántica.
- (P3) Las mentes tienen contenidos; específicamente, tienen contenidos intencionales o semánticos.
- (P4) Los programas son definidos formalmente o sintácticamente.

Conclusiones

- (C1) En sí misma, la instanciación de un programa nunca es suficiente para tener una mente (por P2, P3 y P4).
- (C2) Cualquier artefacto que tenga una mente tendría que tener poderes causales equivalentes (al menos) a los del cerebro (por P1).

Corolarios

- La manera en que el cerebro causa a las mentes no puede ser solo por la instanciación de un programa (por P1 y C1).
- Para cualquier artefacto que tenga una mente, el programa por sí mismo no sería suficiente para proveerle tal mente. El artefacto tendría que tener poderes causales equivalentes al cerebro (por C1 y C2).

El argumento de Searle es, probablemente, el argumento contra el funcionalismo y el concepto de inteligencia artificial que más atención ha recibido, y sin duda ofrece una potente explicación de la diferencia entre los aspectos cualitativos de un estado mental y los de un estado funcional. Las objeciones a dicho argumento son de varios tipos, aquí vamos a atender a dos de ellas porque precisamente podría considerarse que el texto de Dennett “¿Dónde estoy yo?”, que forma parte de la parte práctica de este trabajo, incluye dichas réplicas.

Réplica del sistema. Por una parte, algunos autores consideran que obviamente el sujeto dentro de la habitación no comprende chino, pero que el sistema sí lo hace. Atribuir una mente al sujeto dentro

de la habitación sería el equivalente de atribuirle conciencia a una neurona, o una parte de nuestra red neuronal. En el caso concreto de “¿Dónde estoy yo?”, vemos que el yo entendido como el punto de vista es mayor que la suma de las partes de las que se compone, no es una parte dentro del sistema sino del sistema en su conjunto.

Réplica de la variante robótica. Por otra parte, algunos filósofos critican a Searle porque su experimento no responde a la forma en que efectivamente se aprende un lenguaje, y que si una mente humana debiera manejar un lenguaje de la forma en que se hace en la habitación china, tampoco comprendería lo que dice. Es relevante la experiencia lingüística que supone interacción con el medio para aprender un lenguaje y, por tanto, si la habitación china se materializara en forma de robot con sensores y efectores, y pudiera relacionarse adecuadamente con el entorno físico, entonces dicha variante del programa podría comprender lo que aprendiera (en este caso, chino). Algo así vemos en “¿Dónde estoy yo?”, en que queda más o menos claro que el cuerpo no es solo el cascarón de la mente, pues recibir los inputs a través de un cuerpo y de forma vivencial o experiencial es relevante para la formación misma de una mente.

II.2. La belleza, forma sin función⁷

II.2.1. El juicio de gusto en Kant

La *Crítica del juicio* de Kant se publicó por primera vez en 1790, aunque el filósofo empezó a trabajar en ella ya en 1787, tras la publicación de la segunda edición de la *Crítica de la razón pura*. Dicha obra, supone el cierre del sistema crítico kantiano, Kant señala que “todas las facultades del alma pueden reducirse a tres, que no se dejan deducir de una base común, y son: la *facultad de conocer*, el *sentimiento de placer y dolor*, y la *facultad de desear*.” (Kant 1790, Introducción, III, p. 103) De la primera se ocupa la *Crítica de la razón pura*, de la tercera la *Crítica de la razón práctica*:

Ahora bien: entre la facultad de conocer y la de desear está el sentimiento del placer, así como entre el entendimiento y la razón está el Juicio. Es, pues, de suponer, al menos provisionalmente, que el Juicio encierra igualmente para sí un principio a priori, y que, ya que necesariamente placer y dolor van unidos en la facultad de desear (...), realiza también un tránsito de la facultad pura del conocer, o sea, de la esfera de los conceptos de la naturaleza a la esfera del concepto de libertad, del mismo modo que el uso lógico hace posible el tránsito del entendimiento a la razón. (Kant 1790, Introducción, III, p. 104)

Por tanto, la *Crítica del juicio* va a tratar de mediar entre la necesidad y la libertad, encargándose, en cierto sentido, de suministrar motivos, buenas razones, para admitir los postulados de la razón práctica, gracias al “principio de finalidad de la naturaleza”, un concepto trascendental que no es “ni un concepto de la naturaleza ni un concepto de la libertad [...] sino que representa tan solo la única manera como nosotros hemos de proceder en la reflexión sobre los objetos de la naturaleza, con la intención puesta en una experiencia general y conexas” (Kant 1790, Introducción, V, p. 111). La tercera crítica supone la reconciliación entre la razón pura y la práctica, entre el mundo empírico y la moralidad. La belleza nos deja intuir el orden del mundo, la finalidad de la naturaleza (en un “como si”), esto es, la norma en la contingencia empírica, lo moral en la materia.

No obstante, aquí no vamos a interesarnos más allá de estas consideraciones generales en el lugar de la *Crítica del juicio* en la filosofía kantiana, pues lo que nos interesa es estudiar el concepto de forma, y vamos a hacerlo abordando la teoría kantiana de la belleza. Así, la *Crítica del juicio* se

⁷ Esta parte puede complementarse con el apartado “Una nueva estética” de la unidad de filosofía “Ilustración y salones: nuevos contenidos en un viejo continente” dentro de este mismo proyecto.

compone de dos partes, una dedicada al juicio estético y otra al juicio teleológico, y a su vez, la primera está dividida en dos secciones y la primera de estas en dos libros, el de la analítica de lo bello y el de la analítica de lo sublime. Aquí vamos a interesarnos únicamente por este primer libro.

Al margen del cierre de su sistema, Kant pretende responder a la pregunta estética que también había preocupado a la generación inmediatamente anterior de pensadores, a saber, “¿Cómo se fundamenta la universalidad del juicio de gusto?”. Tras la entrada en crisis de lo que Tatarkiewicz llama la Gran Teoría de la belleza, que entendía esta última como universal, racional y objetiva, los filósofos empiristas, notablemente Hume⁸, tratan de responder a dicha cuestión buscando la universalidad en la configuración del sujeto. Para Hume, de hecho, nuestra igualdad fisiológica y psicológica es la base del juicio de gusto, y dicha igualdad material puede permitirnos aspirar a una comprensión universal del juicio de gusto.

Frente a esta vía psicológica, Kant optará por la vía trascendental: para que el juicio de gusto pueda ser realmente universal, ha de ser necesario, la igualdad de hecho (la material, la psicológica) es *a posteriori* y no *a priori*, luego nunca podría convencernos de su necesidad (en una generalización de ese calibre, respecto a lo bello, ¿quién podría estar seguro de no llegar a conocer a nadie con una configuración psicológica diferente?). Debe basarse pues el juicio de gusto en las facultades.

Según Kant, el juicio de gusto, que define como “la facultad de juzgar lo bello”:

- a) No proporciona conocimiento.
- b) Es desinteresado.
- c) Es universal.

a) ¿Por qué no proporciona conocimiento?

“El objeto es bello” (juicio de gusto), no implica conocimiento acerca del objeto, porque no predica ninguna cualidad del objeto (como serían el peso o el tamaño), la belleza no es un rasgo del objeto. Ahora, tampoco es una proyección subjetiva. El juicio de gusto es *reflexionante* (no transmite conocimiento porque sólo lo particular nos es dado, sin ley) y no *determinante* (el juicio normal, en que lo particular es subsumido bajo lo general que nos viene dado, o por una ley natural para los hechos o por la libertad para las acciones). En el juicio reflexionante lo particular *busca* lo general, pero no se subsume bajo una ley porque no hay leyes en la naturaleza, debe darse a sí mismo una ley, un principio, que es condición de todos los juicios dependientes de la experiencia. No es un juicio empírico, sino su condición. Y puesto que las leyes de la naturaleza se basan así en nuestro entendimiento, las leyes particulares empíricas han de entenderse “como si” hubiese un entendimiento tras ellas. Así, el principio que se da el juicio reflexionante es un principio trascendental (porque no procede de la experiencia, sino que es condición de toda experiencia posible): *el de la finalidad (el concepto) de la naturaleza en su diversidad*.

Esa finalidad formal de la naturaleza es un principio subjetivo del Juicio que no añade a la naturaleza, sino el modo (el “como si”) de entenderla nosotros, de verla general y conexas, de unirla en su variedad. No hay conocimiento exterior que el sujeto pueda poseer, sino que el conocimiento siempre parte del sujeto y es en relación a él: no hay una naturaleza exterior dada, sino fenómenos singulares que responden a ciertas normas, sin un principio que articule esas normas no hay naturaleza como todo, sino mera diversidad natural. Pero ese principio lo pone el ser humano. De la naturaleza recibimos experiencias inconexas, el principio de la finalidad nos la presenta como un

⁸ Véase la nota anterior.

todo ordenado, dicho principio se entrevé en el juicio reflexionante y *su descubrimiento provoca placer.*

b) ¿Por qué es desinteresado?

Porque el placer provocado por el juicio de gusto proviene del descubrimiento de la unión de las leyes empíricas, no de las cualidades del objeto. El juicio de gusto no satisface interés útil alguno. Si lo hace, no es juicio estético sino moral (o moralizante), como cuando se juzga una obra de arte por los efectos piadosos o el ardor revolucionario que provoca en el espectador. El juicio de gusto debería dar placer en sí mismo y no por remitir a un interés ulterior. ¿Qué entiende Kant por **interés**? “Satisfacción que unimos con la representación de la existencia actual o posible de un objeto.” (Kant 1790, Primera Sección, Primer Libro, Primer Momento, § 2, p. 132) Se relaciona pues con la facultad de desear y se corresponde a la satisfacción del deseo.

Dos candidatos a placer desinteresado que sin embargo no son tales son lo agradable y lo bueno. Es cierto que lo agradable parece inmediato y no aporta conocimiento alguno del objeto, sin embargo, en el juicio sobre lo agradable se expresa un interés y se excita el deseo hacia el objeto, y si no es así, entonces no es un juicio. Lo mismo ocurre con lo bueno, en que se transluce un interés por el bien moral, remite pues a fines superiores y no es pues desinteresado. Lo agradable y lo bueno no son pues sino tipos de deseos (aunque muy diferentes).

El juicio de gusto corresponde a una actitud contemplativa que mezcla la inmediatez de lo agradable (pero debe ser juicio y no mera sensación), con el desinterés por la existencia o los valores morales del objeto. ¿Por qué place entonces el juicio estético? Porque excita dos facultades cuyo libre juego place: la imaginación y el entendimiento. A diferencia de su uso en el acto de conocimiento, en que la sensibilidad (intuición sensible) da el contenido, la imaginación lo ordena y el entendimiento lo categoriza, en el juicio estético no hay conocimiento porque precisamente no hay interés por el contenido, y por ello el placer es desinteresado porque surge en el libre juego de imaginación y entendimiento en cuanto que son las facultades de representar. *Se trata de una representación placentera, sin contenido.*

c) ¿Por qué es universal?

Para ser juicio, el juicio de gusto ha de ser universal o de lo contrario sería mera opinión (subjetiva). ¿Cómo entender la universalidad? Lo afirmado ha de proponerse universalmente pero no es necesario que todo el mundo esté de acuerdo en su contenido. Y aunque el juicio estético se formule como un “me parece bello”, tiene pretensión de universalidad y no de simple parecer inconmensurable. Y aquí es donde Kant aborda el problema de la estética principal de la estética, mencionado más arriba: ¿Cómo puede ser universal lo que no ofrece conocimiento objetivo alguno y es desinteresado, o que no provoca acuerdo? ¿No parece esto conducir precisamente a la subjetividad?

La respuesta kantiana es que la universalidad no depende de la suma de los gustos, sino de la necesidad interna del juicio, es trascendental. No depende de los juicios como hechos independientes, sino de su condición de posibilidad. Y esa condición de posibilidad depende de las facultades del conocimiento. El placer estético viene dado por el libre juego de la imaginación y el entendimiento con ocasión de una representación, y ambos son condición universal para el conocimiento de *todos* los individuos. Ese libre juego pasa desapercibido en el conocer algo, pero pasa a primer plano cuando prescindimos del contenido y nos atenemos a la formalidad del proceso (el libre juego de las facultades), cuando la propia naturaleza del objeto (su ausencia de contenido)

nos remite a esa formalidad que contemplamos estéticamente. *Esto es posible para todos, de ahí la universalidad del juicio estético.*

II.2.2. La belleza en Kant

La belleza es la única categoría estética propiamente dicha según Kant, la que se da en los juicios de gusto que, además de causar placer, afirman o niegan la belleza, no como cualidad del objeto, sino como exhibición de su formalidad.

¿Qué es bello entonces?

- Según el juicio de gusto: es bello lo que obliga a reflexionar (contemplar, representar) sobre el libre juego de las facultades porque carece de contenido cognoscible (es bello porque asentimos al placer que nos suscita), por ello es juicio reflexionante. Un juicio que puede ser implícito, es asentimiento por el placer mismo suscitado.
- Según la naturaleza: es bello aquello cuya formalidad cumple en sí la ley de la naturaleza, pues en tanto que pura formalidad es cumplimiento del principio trascendental del juicio (de la finalidad de la naturaleza) que es condición de toda experiencia (lo bello natural es accesorio, y por ello parece que ha sido hecho con desinterés, voluntariamente, para expresar la finalidad misma, la formalidad).

La belleza es pues estrictamente formal. Bello es lo que, sin concepto (luego no constituye conocimiento alguno), es representado como objeto de una satisfacción universal. Universal porque sin tener concepto, lo busca, tiende a él como un fin. Se trata de un pensamiento sin un concepto determinado del entendimiento. En resumen, bello es lo que agrada de acuerdo con el juicio de gusto, lo cual conlleva cuatro características de lo bello (aunque la 2ª y 4ª coinciden) derivadas de las cuatro clases de categorías:

1. Por su cualidad: "Gusto es la facultad de juzgar un objeto o una representación mediante una satisfacción o un descontento, *sin interés alguno*. El objeto de semejante satisfacción llámase *bello*." (Kant 1790, Primera Sección, Primer Libro, Primer Momento, § 5, p. 141)
2. Por su cantidad: "*Bello* es lo que, *sin concepto*, [es representado como objeto que] *place universalmente*." (Kant 1790, Primera Sección, Primer Libro, Primer Momento, § 9, p. 152)
3. Por su relación: "*Belleza* es forma de la *finalidad* de un objeto en cuanto es percibida en él *sin la representación de un fin*." (Kant 1790, Primera Sección, Primer Libro, Primer Momento, § 17, p. 175) Por ello es subjetiva y no una cualidad del objeto, aunque algo de ella dependa del objeto.
4. Por su modalidad: "*Bello* es lo que, sin concepto, es conocido como objeto de una *necesaria satisfacción*." (Kant 1790, Primera Sección, Primer Libro, Primer Momento, § 22, p. 178)

La 1ª remite al desinterés, la 2ª a la universalidad, la 3ª a la ausencia de conocimiento, y la 4ª a la universalidad.

Kant distingue dos tipos de belleza: *libre* (de una relación conceptual) y *adherente* (regida por el concepto). Lo bello libre es aquello bello que no significa nada en sí mismo (los arabescos, las grecas), que no representa nada de lo que tengamos concepto definido. Lo bello adherente presupone un propósito que establece lo que debería ser la cosa (un hombre, un caballo, una iglesia), en referencia a un concepto definido, con lo que no es puro sino dependiente. La belleza libre para Kant es superior (el concepto, el "mensaje", el fin moralizante, limitan el arte y destruyen la belleza).

En general la belleza natural es por ello superior a la artificial, pues el producto del arte está concebido con un fin conceptual (pero no todas las bellezas naturales son libres ni las artificiales adherentes [ver imágenes de arte abstracto]). El arte es tanto más bello cuando parece un producto natural (véase el producto del genio, “la disposición natural del espíritu (*ingenium*) mediante la cual la naturaleza da la regla al arte”) y la naturaleza tanto más bella cuando parece un producto artístico. Esto es porque el arte es bello conforme a un fin y sin embargo provoca sentimiento de libertad en el juego de nuestras facultades de conocer, tiene fin pero parece no haber sido sometido a ese fin.

III. Materiales para el aula

Sesión 1: El problema mente-cerebro

Visión general del problema y las características de lo mental.

Sesión 2: El funcionalismo

Las teorías reduccionistas y la crítica de Putnam a dichas teorías. Se puede abordar directamente algunos de los fragmentos citados del texto de Putnam. El funcionalismo.

Sesión 3: Los problemas del funcionalismo

Las críticas al funcionalismo. Estudio de los distintos argumentos a través de sus esquemas. Discusión, análisis y/o formalización (véase el Anexo). Preparación de la lectura de “¿Dónde estoy yo?” de Daniel Dennett o visionado de su versión en vídeo (véase el Anexo).

Sesión 4: Metodología del comentario de texto y la teoría del gusto de Kant

Se enseña la forma en que se lleva a cabo un comentario de texto filosófico y se explica la teoría del juicio de gusto en Kant a partir del análisis conjunto en clase del texto siguiente:

El juicio de gusto no es un juicio de conocimiento, y por tanto no es científico sino estético, con lo que quiero decir que es un juicio cuyo fundamento solo puede ser subjetivo. [...] Aquí no se indica nada en el objeto, sino que tenemos una sensación de nosotros mismos en cuanto afectados por la representación. [...]

La satisfacción relacionada con nuestra idea de la existencia de un objeto se llama interés. Esta satisfacción, pues, está siempre relacionada con la facultad del deseo, o bien afectándola o como necesariamente relacionada con lo que la afecta. Pero cuando la cuestión es si una cosa es bella, no queremos saber si algo depende o puede depender, para nosotros o para otros, de la existencia de ese objeto, sino solo como lo estimamos en mera contemplación.

Kant, *Crítica del juicio*, Sección 1ª, Libro 1º, § 1 y 2

Sesión 5: La teoría de la belleza de Kant

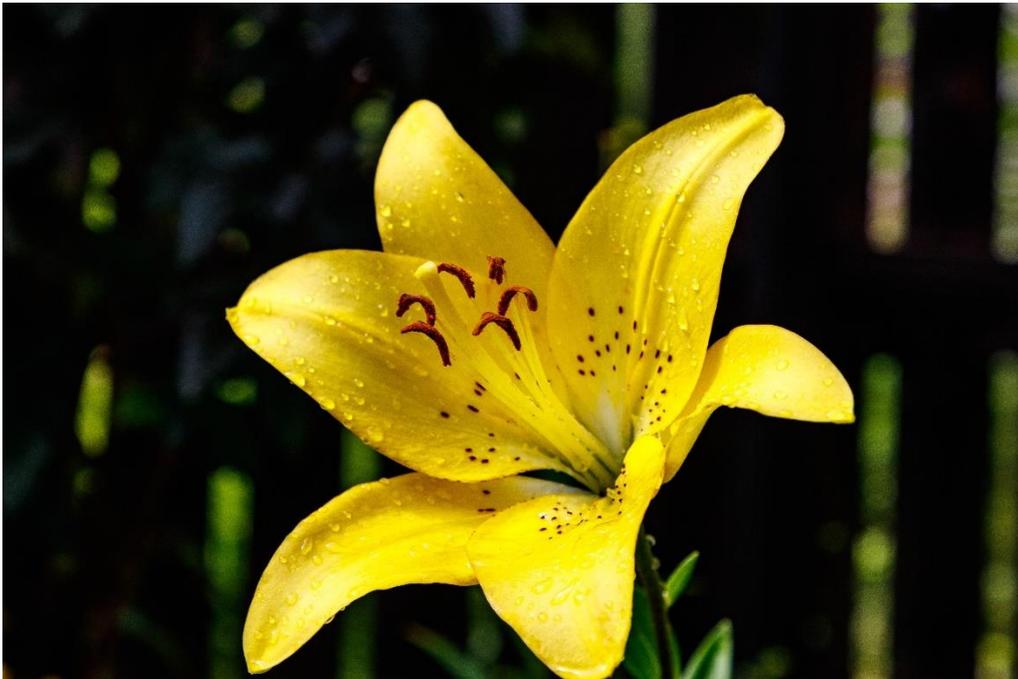
Se propone en primer lugar la lectura del siguiente texto de Kant para explicar la diferencia entre belleza libre y adherente, y tras ello una serie de imágenes con las que ilustrar dicha diferencia a los alumnos:

[E]l juicio de gusto que declara bello un objeto por referencia a un concepto definido no es puro. Hay dos clases diferentes de belleza: la belleza libre (pulchritudo vaga) y la belleza meramente dependiente (pulchritudo adhaerens). Aquella no presupone ningún concepto de lo que debería ser la cosa; esta sí, y presupone también la conformidad de la cosa a ello. [...] Los diseños decorativos griegos, los follajes en márgenes o en papel de pared, y demás, no

significan nada en sí mismos: no representan nada de que tengamos concepto definido y son bellezas libres. También podemos contar como parte de la misma clase lo que en música se llama fantasías, o composiciones sin tema, y de hecho, toda música sin palabras. [...] Pero la belleza de la humanidad, sea hombre, mujer o niño, o de un caballo y un edificio, sea iglesia, palacio, arsenal o casa de verano, presupone un propósito que establece lo que debería ser la cosa (esto es, un concepto de su ideal) y por lo tanto solo es belleza dependiente.

Kant, *Crítica del juicio*, Sección 1ª, Libro 1º, § 16

Belleza libre natural



CC0 Dominio público. Gratis para uso personal y comercial. No se requiere atribución <https://pxhere.com/es/photo/1634925>

Belleza adherente natural



CC0 Dominio público. Gratis para uso personal y comercial. No se requiere atribución <https://pxhere.com/es/photo/946218>

Belleza libre artificial



Detalle de Greca en Mitla, Oaxaca, por Erick Osorio – Obra propia, CC BY-SA 3.0, <https://commons.wikimedia.org/w/index.php?curid=21359549>

Belleza adherente artificial



El descenso de Rogier van der Weyden, imagen de José-Manuel Benito Álvarez, CC BY-SA 2.5, <https://commons.wikimedia.org/w/index.php?curid=4727461>

Belleza adherente artificial



Bolchevique (1920) de De Boris Kustódiev – Obtenido de <http://www.abcgallery.com/K/kustodiyev/kustodiyev15.html> , Dominio público, <https://commons.wikimedia.org/w/index.php?curid=213178>

Comentarios a las imágenes:

La flor puede ser estudiada científicamente, pero es bella en su formalidad ajena a la ley natural, sino al principio trascendental del juicio, que es principio precisamente por imponer formalidad sin contenido. Del mismo modo la greca es belleza libre en tanto no representa nada de lo que tengamos un concepto definido.

Por otra parte, el cuerpo humano, tanto natural, como en su representación pictórica, es belleza adherente, pues place por la representación de su concepto, es deseable (rostro bello) o representa un ideal de bondad (Cristo, realismo socialista).

ANEXO – MATERIALES PARA LAS ACTIVIDADES DE EVALUACIÓN

Filosofía de la mente – Lectura de “¿Dónde estoy yo?” de Daniel Dennett

El texto a leer para reflexionar sobre el problema mente/cuerpo se halla en la compilación de ensayos de 1981 *Brainstorms: Philosophical essays on mind and psychology*, publicado en Cambridge (Mass.) por MIT Press. Dicho libro no ha sido publicado en español, pero puede encontrarse una traducción del texto en este enlace: <https://fdocuments.es/document/donde-estoy-yo-dennett.html> A su vez, dicho relato se convirtió en ficción televisiva, y puede encontrarse dicha versión aquí: <https://www.youtube.com/watch?v=KP7rTp2vwTs> Para trabajar dicho relato, así como el resto de contenidos de filosofía de la mente y del funcionalismo se proponen las cuestiones siguientes para los alumnos (véase la página siguiente):

1/ ¿Quiénes son Yorick, Hamlet, Hubert y Fortinbras? (1 punto)

2/ ¿Es Dennett, Hamlet? ¿Por qué? (1 punto)

3/ ¿Es Dennett, Yorick? ¿Por qué? (1 punto)

4/ ¿De acuerdo con el funcionalismo, puede la suma de Hubert y Fortinbras ser Dennett? ¿Y según el relato? ¿Por qué? (1 punto)

5/ ¿Significan las siguientes palabras de Dennett que la mente podría existir sin un soporte físico? Justifica tu respuesta. (2 puntos)

Pero en mis actuales circunstancias parecía que [...] el lugar donde se encuentra una persona (y, en consecuencia, donde sus pensamientos se ubican, si lo analizamos semánticamente) no es necesariamente donde su cerebro, el asentamiento físico de su «alma», se encuentre.”

6/ ¿Qué quiere decir Dennett cuando afirma lo siguiente? Explica qué ha ocurrido en este punto y qué es lo que se ha probado. Relaciónalo con la crítica al reduccionismo desde la perspectiva funcionalista. (2 puntos)

Se me ocurrió entonces, con una de esas revelaciones de las que uno debería sospechar, que había tropezado con una impresionante demostración de la inmaterialidad del alma basada en principios y premisas fiscalistas. Pues, ¿acaso no me había desplazado a la velocidad de la luz desde Tulsa a Houston al desvanecerse la señal de radio entre Tulsa y Houston?

7/ Hacia el final del relato Dennett considera la posibilidad que aparece en la cita siguiente. Explica y justifica lo que afirma al final de la cita. ¿Cuántas personas existen al final del relato? (2 puntos)

El único aspecto descorazonador del nuevo desarrollo de los acontecimientos era la perspectiva, que no tardé en tener en cuenta, de que alguien desconectara el cerebro de repuesto — Hubert o Yorick, el que fuese— de Fortinbras y lo conectase a un nuevo cuerpo —un recién llegado Rosencratz o Guildenstern—. Entonces (y no antes) sí habría dos personas, eso estaba claro.

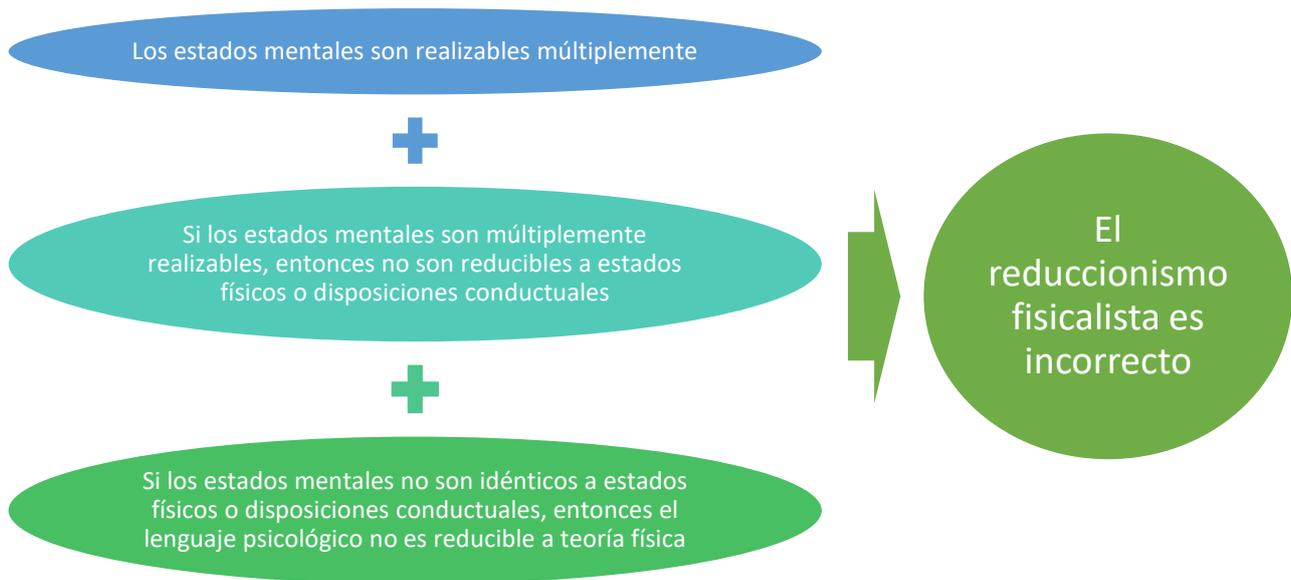
También, si se aborda la lectura del texto tras haberse estudiado el experimento mental de la habitación china de John Searle contra el funcionalismo, se puede instar a los alumnos a que expliquen de qué forma el relato de Dennett supone una refutación del argumento de Searle. O más concretamente, que especifiquen de qué dos formas el relato de Dennett contradice el experimento de la habitación china (pues en él aparecen claramente las réplicas del sistema y de la variante robótica).

Filosofía de la mente – Argumentos para analizar

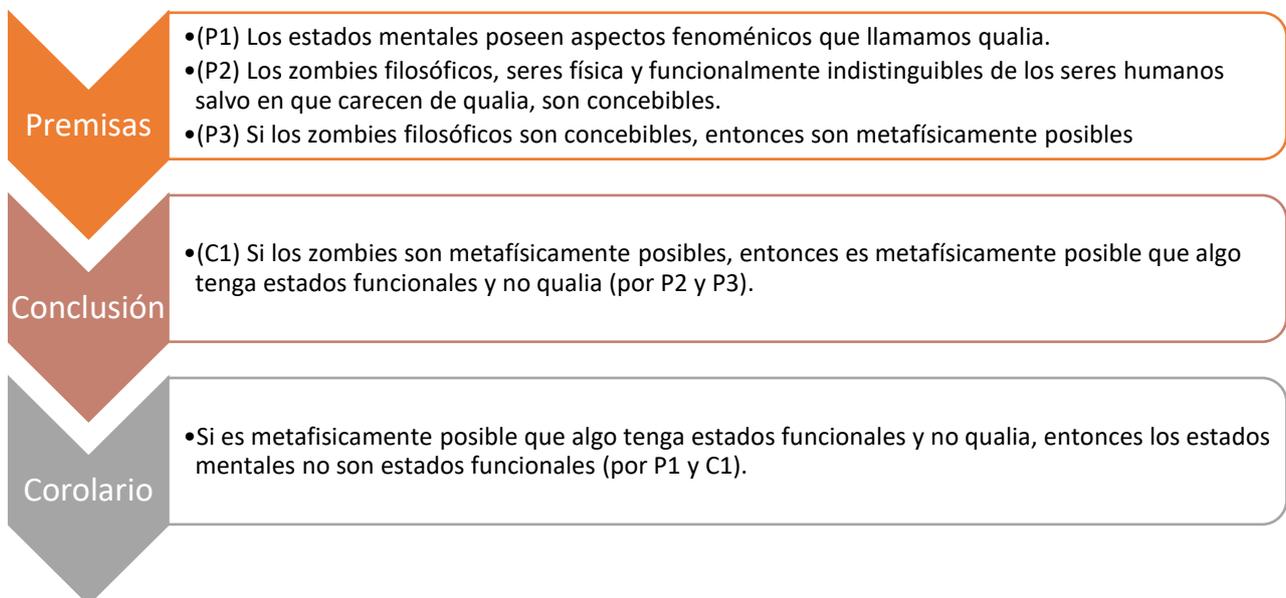
Para trabajar no solo el concepto de función, sino también el de forma, si los alumnos han estudiado ya en el curso lógica formal, se ofrecen claramente contruidos algunos argumentos clásicos del debate sobre el funcionalismo en filosofía de la mente, para que los alumnos puedan formalizarlos y probar su validez, bien valiéndose de tablas de verdad, bien mediante deducción natural.

También, por el hecho de estar expuestos de forma sistemática y clara, pueden tratarse desde la lógica informal y debatirse no tanto atendiendo a su validez como a la verdad de sus premisas, lo cual daría pie a interesantes debates en el aula acerca de muchos de los conceptos fundamentales de filosofía de la mente tales como el de “realización”, “reducción”, “qualia”, “estado mental”, “estado funcional”, etc...

1) *Argumento de la realizabilidad múltiple contra la teoría de la identidad y el conductismo*



2) *Argumento de los zombies filosóficos contra el fisicalismo y el funcionalismo*



3) *Argumento contra la existencia de qualia (o los argumentos anti-fisicalistas y anti-funcionalistas basados en los qualia)*



4) *Argumento de la nación china por el excesivo liberalismo del funcionalismo*



5) Argumento de la habitación china contra el funcionalismo y la IA fuerte

Premisas

- (P1) Los cerebros causan mentes.
- (P2) La sintaxis no es suficiente para la semántica.
- (P3) Las mentes tienen contenidos; específicamente, tienen contenidos intencionales o semánticos.
- (P4) Los programas son definidos formalmente o sintácticamente.

Conclusiones

- (C1) En sí misma, la instanciación de un programa nunca es suficiente para tener una mente (por P2, P3 y P4).
- (C2) Cualquier artefacto que tenga una mente tendría que tener poderes causales equivalentes (al menos) a los del cerebro (por P1).

Corolarios

- La manera en que el cerebro causa a las mentes no puede ser solo por la instanciación de un programa (por P1 y C1).
- Para cualquier artefacto que tenga una mente, el programa por sí mismo no sería suficiente para proveerle tal mente. El artefacto tendría que tener poderes causales equivalentes al cerebro (por C1 y C2).

Teoría estética kantiana – Comentario de texto

Para estudiar el concepto de forma a partir de la estética kantiana, de su teoría del juicio de gusto y su teoría de la belleza, se proponen los siguientes textos (selección extraída de la *Breve historia y antología de la estética* de Jose María Valverde) para hacer su comentario (y a continuación se ofrece una metodología del comentario de texto y una plantilla de corrección):

1) Texto 1

Lo bello es lo que se piensa como objeto de una satisfacción universal aparte de ningún concepto. [...] Una cosa de la que cada cual reconoce que su satisfacción en su belleza es desinteresada, debe ser estimada por cada cual como motivo de satisfacción para todos los hombres. [...]

Una pretensión de universalidad subjetiva debe estar implicada en el juicio de gusto. [...] Uno dice: Esa cosa es bella; no espera meramente el asentimiento de los demás a su juicio de satisfacción porque han estado de acuerdo con él antes: lo exige de ellos. Les critica si juzgan de otro modo y les niega el gusto. [...]

La universalidad de nuestra satisfacción está representada en el juicio de gusto solo como subjetiva. Si estimamos los objetos meramente por conceptos, se pierde toda idea de belleza. Así que no puede haber regla por la cual nadie esté obligado a reconocer algo como bello. Nadie permite que su juicio sobre la belleza de una casaca, de una casa, de una flor, esté coaccionado por razones y principios. Quiere tener la cosa delante como si su satisfacción fuera sensorial, pero si la llama bella, pretende tener de su parte la opinión universal, mientras que la sensación es privada y no decide nada más allá de la satisfacción del que la tiene. [...]

Kant, *Crítica del juicio*, Sección 1ª, Libro 1º, § 6

2) Texto 2

Las facultades de conocimiento puestas en juego por la idea están en tal caso en juego libre, porque ninguna concepción definida de la naturaleza del objeto las confina a un principio particular de conocimiento. Así el estado de la mente al tener tal idea debe ser un sentimiento del libre juego de nuestras facultades al usar una percepción para propósitos de conocimiento en general. [...] Esta estimación meramente subjetiva (estética) del objeto, o de nuestra idea de él, precede a nuestro placer en él, y ocasiona este placer en la armonía de nuestras facultades de conocimiento. [...] En un juicio de gusto atribuimos como necesario a todo el mundo el placer que sentimos nosotros; como si la belleza que adscribimos a una cosa hubiera de ser considerada una propiedad que se sigue de su concepción; aunque la belleza, aparte de la relación con nuestro sentimiento, no es nada en sí misma. [...]

Un propósito, considerado como causa de nuestro placer, implica un interés que ocasiona nuestro juicio sobre el objeto agradable. Entonces, no puede ser ningún propósito nuestro lo que ocasione el juicio de gusto.

Kant, *Crítica del juicio*, Sección 1ª, Libro 1º, § 9

3) Texto 3

[U]n mero color, por ejemplo el verde de un césped, y un mero tono (en cuanto diferente de sonido y ruido), por ejemplo, de un violín, son llamados bellos en sí mismos por la mayoría de la gente, aunque ambas cosas parecen depender de la mera materia de nuestras ideas, esto es, simples sensaciones, y por tanto solo merecerían ser llamadas agradables o encantadoras. Sin embargo, podemos observar aquí que las sensaciones de color y tono pretenden ser contadas como bellas, en la medida, y solo en la medida, en que son puras. Pues la pureza es un carácter de su forma y también el único carácter de esas ideas que puede ser estimado universalmente. [...]

Kant, *Crítica del juicio*, Sección 1ª, Libro 1º, § 14

Metodología del comentario de texto

El comentario de texto consiste fundamentalmente en extraer el interés filosófico de un texto analizándolo de forma ordenada. En este esquema he separado estas dos tareas, pero pueden llevarse a cabo de forma conjunta.

Introducción

- *Presenta* brevemente el texto y su autor (pero no recites su biografía).
- Indica cuál es el *problema general* al que trata de dar respuesta el texto.
- Señala cuál es la *tesis* defendida por el autor con respecto a ese tema.
- Haz un esquema con las *partes principales de la argumentación* (no se trata de hacer un esquema con gráficos y flechas, sino bien redactado, como una pequeña sinopsis del texto en el que destaquen los diferentes argumentos pero sin entrar en detalle).
- Anuncia el *plano de la argumentación* que vamos a seguir (ej: “En un primero momento nos ocuparemos de... más adelante... por fin, en un tercer y último apartado... lo cual nos permitirá concluir...”).

Desarrollo

1. Análisis

En este apartado se debe mostrar *la lógica del texto* y explicar *sus ideas principales*.

El texto propuesto ha sido escogido de tal forma que permite plantear, véase resolver, un *problema filosófico*. Se trata precisamente de abordar indirectamente ese problema a través del enfoque que le da el texto.

Por norma general el autor del texto defiende una determinada postura con respecto a ese problema, y lo hace a través de una *argumentación* que tendremos que desentrañar. No se trata de yuxtaponer consideraciones fragmentarias en torno al texto, sino de darle un sentido a partir del problema principal que plantea el texto. Se trata de mostrar qué es lo que está en juego y cómo lo aborda el autor, qué hace con ello. Se trata de mostrar la *coherencia interna* del texto (no forzosamente siguiendo el orden del texto).

Un problema en esta parte es parafrasear el texto (repetir el texto con otras palabras). Para prevenir este problema habrá que *definir* los conceptos que aparecen en el texto, *analizar* cada noción importante, *mostrar los nexos lógicos* que llevan de una idea a otra, *explicitar* lo implícito y *enriquecer* el texto con nuestra propia cultura filosófica. Otro gran problema sería convertir el comentario en un recital de nuestros conocimientos acerca del autor o su filosofía, esto es estar fuera del tema.

2. Discusión/perspectiva filosófica del texto

En esta parte *reaccionamos* brevemente ante el texto.

Debemos tomar un poco de distancia con respecto al texto para mostrar las *perspectivas* que abre, desarrollar sus consecuencias filosóficas (lo que implica o significa), *confrontando la tesis del autor o su argumento a perspectivas contrarias o contraargumentos*. Si tenemos algunos conocimientos de historia del pensamiento nos pueden ser útiles aquí siempre que no caigamos en el error del recital de conocimientos señalado más arriba.

Debemos también aquí, por fin, *tomar partido* con respecto al texto, criticar, para bien o para mal, la tesis o el argumento del autor, y definir nuestra postura con respecto al problema filosófico abordado por el texto. Esta postura no tiene por qué ser concluyente, se trata de dar razones y no de dar opiniones.

Fundamentalmente hemos de *completar* el texto con algunos aspectos del problema que consideremos relevantes pero que han sido dejados de lado por el autor, mostrar los *límites* de la perspectiva del texto, formular con prudencia y siempre justificándolas nuestras propias *objeciones*, mencionar *sistemas filosóficos contrarios* pero siempre sin olvidar mencionar la *pertinencia* y el *interés* del punto de vista introducido por el autor.

El gran problema a evitar en esta parte del comentario es la de convertir el texto en un "pretexto" para exponer nuestras ideas filosóficas. No debemos convertir el comentario de texto en una disertación.

Conclusión

- Haz una *recapitulación* del desarrollo en unas pocas líneas.
- Señala, sin dogmatismos y sin sentencias lapidarias, si el autor del texto logra alcanzar en el mismo el *objetivo* que se había impuesto. No añadas argumentos o ideas que no se abordaron antes.
- Evita dar tu opinión en dos líneas al final del texto, nuestra opinión personal debe aparecer en la forma en que hayamos abordado el pensamiento del autor tratado, debe ser el resultado lógico de lo planteado en el desarrollo.

Plantilla para la corrección del comentario de texto

Introducción - 2				
El alumno presenta el texto y el autor . (0,5)				
El alumno especifica cuál es el problema del texto y lo explica brevemente, demostrando comprender su alcance. (0,5)				
El alumno especifica la tesis del autor y la explica brevemente. (0,5)				
El alumno introduce el desarrollo anunciando brevemente el contenido de su análisis y discusión . (0,5)				
Desarrollo - 7				
El alumno identifica las ideas principales del texto sin repetirlas literalmente y se refiere a ellas citándolas de manera adecuada. (1,25) (1 en textos breves)				
El alumno explica adecuadamente el significado de las ideas del texto con claridad y precisión. (1,25) (1 en textos breves)				
El alumno emplea con precisión los conceptos clave del texto , es capaz de definirlos, explicarlos o emplearlos correctamente. (1)				
El alumno explica la relación existente entre las distintas ideas del texto, y reconstruye su estructura argumentativa especificando cómo se sustenta la tesis y resumiendo el proceso. (1,5) (1 en textos breves)				
El alumno comprende el problema discutido en el texto y explica las consecuencias de defender la respuesta del autor y los principales contraargumentos. (1,5) (3 en textos breves)				
La composición del alumno es rica, este escribe de forma correcta no solo en lo que concierne a la sintaxis y la ortografía, sino también al estilo. (0,5)				
Conclusión - 1				
El alumno recapitula los aspectos principales del texto, cómo ha abordado el autor ese problema filosófico y evalúa su respuesta. (1)				
Calificación				

Bibliografía:

- BLOCK, Ned (1978), "Troubles with functionalism", traducción "Las dificultades del funcionalismo" en Rabosi (1995).
- BOZAL, Valeriano (ed.) (1996), *Historia de las ideas estéticas y de las teorías artísticas contemporáneas. Volumen I*. Madrid: Visor.
- DENNETT, Daniel (1981), "Where am I?" en *Brainstorms: Philosophical essays on mind and psychology*. Cambridge (Mass.): MIT Press. Puede encontrarse una traducción del texto en este enlace: <https://fdocuments.es/document/donde-estoy-yo-dennett.html>
- _____ (1995), *La conciencia explicada*. Barcelona: Paidós.
- JAWORSKI, William (2011), *Philosophy of mind: a comprehensive introduction*. Oxford: Wiley-Blackwell.
- KANT, Immanuel (1790, edición de 1997), *Crítica del juicio*. Madrid: Editorial Espasa Calpe.
- MOYA, Carlos (2006), *Filosofía de la mente*. Valencia: Publicaciones de la Universidad de Valencia.
- PUTNAM, Hilary (1967) (Trad. 1981 por VALDÉS, Margarita), *La naturaleza de los estados mentales*. México: Instituto de Investigaciones Filosóficas. Disponible en el enlace <https://www.icesi.edu.co/blogs/experimentosmentales/files/2010/05/LA-NATURALEZA-DE-LOS-ESTADOS-MENTALES2.pdf>
- RABOSI, Eduardo (ed.) (1995), *Filosofía de la mente y ciencia cognitiva*. Barcelona: Paidós.
- SEARLE, John (1989), "Minds and brains without programs", traducción "Mentes y cerebros sin programas" en Rabosi (1995) .
- _____ (2001), *Mentes, cerebros y ciencia*. Madrid: Cátedra.
- TATARKIEWICZ, Wladislaw (2002), *Historia de seis ideas: Arte, belleza, forma, creatividad, mimesis, experiencia estética*. Madrid: Editorial Tecnos
- TURING, Alan (1950), "Computing machines and intelligence", *Mind* **LIX** (236): 433-460. Puede accederse a una traducción al español aquí: <http://xamanek.izt.uam.mx/map/cursos/Turing-Pensar.pdf>
- VVAA (1990). *Estudios sobre la "Crítica del juicio"*. Madrid: Antonio Machado.
- VALVERDE, José María (1987), *Breve historia y antología de la estética*, Barcelona: Editorial Ariel.