

El problema de la conciencia: de lo individual a lo social.

Domingo Fernandez Agis. Universidad de la Laguna.

Resumen

Dos sentidos del término conciencia han convivido desde que tenemos constancia del uso del mismo: el psíquico y el moral. Un rasgo distintivo del pensamiento moderno consiste en considerar que el primero ha de prevalecer sobre el segundo. Sin embargo, la propia evolución del concepto parece exigir en nuestros días una apelación a lo social para completar su determinación.

La conciencia como realidad psíquica

En fecha tan temprana como 1538 establecía Luis Vives, en su *Tratado del alma*, una nítida separación entre el ser humano y los animales basándose, no tanto en la presencia o no de conciencia, cuanto en el contenido de la misma. Estimaba que la conciencia humana aparece ligada a la capacidad de *recordar* y *planificar*. Esos dos polos, por medio de los cuales podemos apreciar cómo la mirada humana se hace visión simbólica, apuntan a que el lenguaje responde a nuestra necesidad de trascender lo meramente representado a partir de los datos sensibles, o tal cómo él nos dice, los frutos de la *fantasía*. Es obligado, aunque para algún lector resulte ocioso, recordar que la *fantasía* es en esa época considerada como la facultad de construir perceptos, a partir de la información sensible recogida por nuestros órganos receptores en el entorno inmediato. Por tanto, el reto al que ha de responder la conciencia humana se define por la construcción y el empleo continuado de un sistema simbólico que nos permita afianzar cierta estabilidad, más allá de la pluralidad irreductible propia de los datos de la experiencia sensorial. Se trata, por tanto, de la construcción de un mundo donde antes sólo había diversidad, de fundamentar la posibilidad de una unidad coherente allí donde únicamente parecía existir la heterogeneidad contradictoria de lo concreto.

Para ello, como nos dice Vives, “no se inclina la aspiración entera de la mente ante la utilidad o daño presente sino que recuerda lo pasado y conjetura acerca de lo futuro; busca el juicio de lo verdadero y de lo falso, cosa de que no se preocupa el animal, sólo atento a mirar lo que conviene al cuerpo y lo que le daña, sin más que el arrebató de la fantasía.

De esa potencia de la mente trae su fuente y origen el lenguaje, expresión de cuanto en ella se contiene y facultad de que asimismo carecen los brutos” (Vives, 1538: 57). Con ese modo de representación simbólica nos distanciamos de forma definitiva del nivel en que se encuentran el resto de los animales, por más que en el caso de algunas especies haya que advertir la presencia de formas elementales de conciencia.

La originalidad de estos planteamientos basta para concederle a la obra mencionada el valor de encrucijada en la cultura científica y filosófica europea del momento. En el punto de encuentro que ella representa convergen los caminos que vienen de la psicología y la medicina grecolatinas, revitalizadas por el humanismo renacentista, y comienzan las nuevas sendas que culminarán en el proyecto moderno de construir un conocimiento científico completo de la naturaleza humana. Hay que subrayar de la forma más enfática posible cómo aparecen ya en Vives algunas de las claves del enfoque que la Psicología actual hace de la cuestión. En efecto, él es el primero en plantear una estrategia de aproximación que hoy nos resulta común. En este sentido, nos dice que la conciencia “habrá de ser observada en sus operaciones, porque no se ofrece a nuestros sentidos; mientras que con estos, así internos como externos, podemos conocer sus obras” (Vives, 1538: 61).

Por tanto, la conciencia, como cualquier otro fenómeno de nuestro psiquismo, ha de ser inferida a partir de la toma en consideración de sus manifestaciones externas, en la conducta del individuo. Pero no es esto sólo lo que el sabio renacentista sostiene al respecto. Con una rotundidad y claridad admirables, si tenemos en cuenta la época en que estas apreciaciones se formulan, enfatiza Vives que “no nos importa saber qué es el alma, aunque sí, y en gran manera, saber cuáles son sus operaciones” (Vives, 1538: 61).

Por todo ello, es conveniente dejar aquí bien claro que su valerosa toma de posición le sitúa ya en el camino que llevará, siglos después, a la constitución de la psicología como

disciplina científica. Porque Vives renuncia a la especulación sin fundamento y vincula de manera explícita sus objetivos al análisis de todos aquellos datos de los que quepa la posibilidad de establecer un contraste y un conocimiento precisos. Advierte, no obstante, que ese conocimiento no es obtenible de una forma sencilla. Distintos niveles de complejidad nos aguardan a medida que nos adentramos en el mismo. En primer término, en nuestras percepciones se funden y reordenan datos que provienen de diferentes sentidos. El objeto percibido es, por tanto, un constructo mental fruto de la combinación coherente de distintos factores estimativos. Por otro lado, más allá de la construcción misma del percepto, un *sentido interior*, que hemos de identificar con la conciencia, se encarga de vincular unos perceptos con otros, dando de esa manera una utilidad inmediata al contenido de la memoria y posibilitando su articulación con las situaciones presentes. A su juicio, “la distinción y separación de los elementos deben partir de los datos sensibles, pues toda facultad se refiere a aquello en lo cual se ejercita. Aunque existen algunos objetos comunes a varios sentidos, como el movimiento, el tamaño, el número, la forma o figura, el sitio o la posición en un lugar, así como las cosas en ellos comprendidas, que son próximamente comunes a la vista y al tacto, de ningún modo pertenece a los sentidos lo que la facultad interior saca de los conocimientos dados por ellos, verbigracia, la hermosura en la forma o la fealdad, la semejanza y la desemejanza” (Vives, 1538: 44).

La idea funcionalista aparece prefigurada en la analogía que establece entre los procesos intelectivos y los físicos. El organismo humano es una totalidad que va constituyéndose por medio de distintos niveles de integración. De la misma forma que el aparato digestivo integra un complejo y variado conjunto de órganos que son necesarios para realizar las funciones digestivas, el psiquismo humano admite una consideración globalizadora porque es el resultado de una integración análoga. La imaginación, en su acepción ya especificada, la memoria y la inteligencia armonizan sus respectivos funcionamientos gracias a la conciencia, haciendo de esta forma posible establecer la unidad que da sentido a la subjetividad humana. Vives nos dice que, “así como en las funciones de nutrición reconocemos que hay órganos para recibir los alimentos, para contenerlos, elaborarlos y para distribuirlos y aplicarlos, así también en el alma, tanto del hombre como de los animales, hay una facultad para recibir las

impresiones de los sentidos, la cual se llama imaginativa; otra para retenerlas, o sea la memoria; otra que las perfecciona, la fantasía, y finalmente, la que las clasifica según su asenso o disenso, que es la estimativa” (Vives, 1538: 51).

Quizá no haya escapado al lector el carácter prevalente que Vives atribuye a la inteligencia, interpretada ahora como facultad de producir juicios o racionios. La cúspide de la actividad mental humana se concibe y explica desde una perspectiva funcional. Para ello, Vives se propone analizar su funcionamiento específico, así como las relaciones concretas que la inteligencia guarda con otras facultades psíquicas. Para él, el resultado de los juicios que el intelecto elabora es permitir la integración de manera eficiente en la realidad que nos rodea. Tampoco habrá escapado al lector la originalidad de su planteamiento, por lo que se refiere a la visión de las facultades psíquicas, en las que establece una gradualidad en el curso de la comparación entre los animales y el hombre. En éste último, la facultad estimativa procede mediante la construcción de juicios, de pensamiento racional, mientras que en el animal prevalecen las reacciones instintivas y las emocionales, en el sentido más primario del término. Como dice el propio Vives, “la facultad estimativa es aquella que partiendo de las impresiones sensibles produce el acto del juicio, dirigido a distinguir lo que puede ser provechoso o nocivo, puesto que la naturaleza la creó para nuestra salud, y para el conocimiento o estímulo de los sentidos” (Vives, 1538: 53).

Continuando con estas consideraciones iniciales, en las que tan sólo se pretende dejar constancia de la aparición del enfoque moderno del problema de la conciencia, no podemos dejar de señalar, por último, que Vives pretende en su obra esbozar un cierto mapa cerebral que le permita localizar las funciones psíquicas básicas en el lugar físico en que sus operaciones se realizan. En este campo es tal vez donde puede apreciarse de una forma más evidente la radical modernidad de sus planteamientos. Su enfoque constituye un modelo, por primario que éste sea, de cómo ha de combinarse reflexión e investigación empírica. Para él, sin esta última no alcanzaremos jamás un resultado positivo. Resolver el problema de la conciencia nos lleva, de manera inexorable, a enfrentarnos a la cuestión de las relaciones entre la mente y el cuerpo. El abordaje más coherente que pueda hacerse de este problema, una de las grandes cuestiones que la

humanidad se ha planteado y que, como es bien sabido, aún no ha alcanzado a resolver, nos lleva a buscar la ubicación física de las facultades psíquicas. ¿Dónde se acometen, en suma, las funciones básicas que nuestro cerebro tiene que desempeñar?

Se trata, como decimos, de una cuestión fundamental, pues la respuesta a la misma puede darnos la clave para responder de forma coherente al problema mente/cuerpo. Vives sostiene que “otorgó la Naturaleza a estas facultades diversos instrumentos, a modo de laboratorios distintos en las partes del cerebro: en la región anterior de éste se afirma que está el origen y asiento de los sentidos, y que en ese mismo sitio se forma la imaginación; en el centro se hallan la fantasía y la facultad estimativa; la memoria en el occipucio” (Vives, 1538: 54). Parece, por tanto, que Vives traza su esbozo de mapa cerebral siguiendo el criterio de la proximidad relativa al sentido predominante en los humanos adultos, la vista. De forma semejante a como unos mecanismos presuponen el funcionamiento de otros, las facultades psíquicas superiores se apoyan en las inferiores y las estructuras correspondientes se sitúan en consecuencia tras ellas. Hasta llegar a la memoria, que recoge la información pertinente que proviene de cada una de las anteriores. De acuerdo con su crucial importancia, la facultad estimativa que se alimenta de todo lo anterior, coordina y da las respuestas pertinentes en cada situación y se sitúa en el centro. Esta facultad está vinculada de forma directa con la conciencia que, por ello, es ubicada en la misma posición central.

Naturalmente, lo interesante aquí no es discutir el acierto o no de semejante distribución espacial de funciones. No hay que olvidar que estamos hablando de una teoría elaborada en el siglo XVI y es imprescindible tener presentes las limitaciones y restricciones con las que este tipo de investigación podía encontrarse en esa época. Quizá lo que más deba importarnos a este respecto sea, en realidad, el intento de correlacionar la conciencia con la ubicación de cada una de las facultades psíquicas, de acuerdo con un canon explicativo presupuesto. Lo relevante es, en suma, que Vives se plantea que existe un soporte físico de las actividades mentales, que ese soporte es específico y diferenciado para cada una de ellas y, por último, que es un planteamiento coherente tratar de localizar la sede de esas funciones en distintas estructuras cerebrales.

Abundando en este planteamiento, es relevante analizar la relación entre la facultad estimativa y la conciencia, así como sus vinculaciones con lo que Vives denomina fantasía e imaginación, que, como ya se ha dicho, poco tienen que ver con lo que la psicología popular entiende hoy por tales términos.

Tras el cartesianismo y su influencia en la visión de la lógica y sus relaciones con el lenguaje en la Escuela de Port-Royal, caracterizada como es bien conocido por conceder una prioridad incuestionable al pensamiento y considerar que a la base de éste se encuentran ciertos elementos de carácter innato, el planteamiento filosófico del problema de la conciencia experimenta un giro decisivo con la *Crítica de la Razón Pura* kantiana. En esta obra el análisis de los fenómenos relacionados con la conciencia se plantea desde una perspectiva muy diferente a la que Vives proponía. En efecto, para Kant es imposible otro conocimiento de la conciencia que aquel que puede obtenerse de la determinación de ser consciente en el tiempo. Es la temporalidad lo que nos permite una intuición de la existencia de la conciencia, aunque no un conocimiento de la conciencia como objeto, pues de la conciencia en sí misma no puede haber representación alguna. Para el pensador alemán, “el sentido interno por medio del cual el psiquismo se intuye a sí mismo o su estado interno no suministra intuición alguna del alma misma como objeto. Sin embargo, hay sólo una forma determinada bajo la que es posible la intuición de un estado interno, de modo que todo cuanto pertenece a las determinaciones internas es representado en relaciones de tiempo” (Kant, KvR, A23).

No cabe, con respecto a la conciencia, el mismo tipo de conocimiento que se deriva de los juicios cognoscitivos. Éstos tienen, según Kant, una base empírica y un fundamento interno. Por lo que respecta a la primera, el juicio se refiere siempre a algo exterior al sujeto, que puede ser objeto de constatación o comprobación experiencial por parte de otros sujetos. Desde este punto de vista, el juicio nos lleva a advertir que es posible la objetividad del conocimiento, entendida como intersubjetividad de la experiencia. Pero, como decíamos antes, el juicio posee también un fundamento interno que Kant identifica con las categorías del entendimiento, verdaderas estructuras formales inscritas de forma innata en él y que explican el valor universal de la verdad. Bien es cierto que todo esto se refiere, como se ha apuntado, a los juicios cognoscitivos, que son los que nos proporcionan un saber acerca de la experiencia, y no a los juicios reflexivos, que tan

sólo nos permiten atribuir cualidades como las éticas o las estéticas, que no pueden ser objeto de demostración concluyente. Porque, para Kant “la verdad descansa en la concordancia con el objeto y, consiguientemente, los juicios formulados por cualquier entendimiento y relativos a dicho objeto tienen que coincidir (*consentientia uni tertio, consentiunt inter se*). El criterio para saber si el tener por verdadero es convicción o mera persuasión es, por tanto, externo, y consiste en la posibilidad de comunicarlo y en comprobar su validez para toda razón humana. En este último caso parece al menos que la causa de la concordancia de todos los juicios residirá, a pesar de la diversidad de los sujetos entre sí, en la comunidad del fundamento, es decir, en el objeto; al concordar todos en él, queda demostrada la verdad del juicio” (Kant, *KvR*, A820/B848).

Partiendo de estas bases, Kant alude, en primer término, a la posibilidad de expresar lo pensado recurriendo a los símbolos propios del lenguaje. Como segundo criterio, nos orienta hacia el objeto, señalando que el consenso colectivo en torno a él puede considerarse como prueba de la certeza de nuestra representación del mismo. Pero, como bien sabemos, en el análisis del proceso cognitivo que había realizado en las dos primeras partes de esta misma obra, Kant se esfuerza en superar la antinomia existente entre las teorías del conocimiento empirista y racionalista, contemplando éste como resultado de un proceso en el que se combinan elementos estructurales innatos, con la información recogida a partir de la experiencia sensible. Por consiguiente, el propio objeto del conocimiento ha de ser contemplado, no como algo simplemente dado, sino como el resultado de una construcción en la que lo aportado por las estructuras cognitivas del sujeto resulta fundamental. Si la expresión lingüística tiene sentido es porque su fundamento cognitivo recae en esa conjunción de lo innato y lo adquirido característica de la epistemología kantiana.

La inflexión hegeliana

Hasta aquí, hemos recorrido algunas de las etapas de ese difícil camino mediante el que el planteamiento del problema de la conciencia se desprende de las ataduras que lo vinculaban a la visión tradicional, que hacía de ella algo tan próximo como ignoto, el soporte de la moral en su dimensión individual a la vez que una entidad imposible de

ser captada de modo objetivo. Veamos ahora otra inflexión en ese mismo recorrido, en ella parece abandonarse por primera vez la insistencia en la escisión y vuelven a confluír ambos extremos, mostrándonos la conveniencia de subrayar la conexión entre los mismos.

El sistema hegeliano se apoya en la idea de una realidad absoluta que no está más allá de nuestras posibilidades de conocimiento, sino de la que es posible alcanzar un saber, puesto que de alguna forma se halla incurso en aquello que parecía oponérsele. El espíritu, tal como lo denomina Hegel, cuya presencia puede ser reconocida en todo lo existente, no se reconoce de antemano a sí mismo en lo que en concreto existe, sino que se pierde inicialmente en ello, se aliena en una realidad con la que aún no puede reconciliarse. Esa reconciliación será el resultado de un proceso, por medio del cual el espíritu va reconociendo su presencia en el interior de aquello que en principio parecía serle por completo ajeno.

Se ha dicho que su filosofía es más bien una filosofía del concepto que de la conciencia, pero no hay que olvidar que sin la conciencia no se podría superar la alienación del espíritu ni éste podría devenir jamás concepto, es decir, objeto de un conocimiento cierto para nosotros. En cuanto al método que Hegel propone para alcanzar ese objetivo, es mucho lo que se ha dicho hasta ahora acerca de él. En todo caso bastaría para caracterizarlo, al menos en una primera aproximación, afirmar que la dialéctica, más que un simple método de conocimiento, se presenta como una clave de interpretación que se desprende de la propia vida o del movimiento mismo del contenido.

Ortega y Gasset definió con gran claridad las líneas más revelantes de la reflexión en torno al método, al señalar cómo éste prefigura cuál ha de ser el comportamiento de la mente, aún antes de su contacto real con los objetos concretos. Cómo él mismo solía recordar, esto supone una predeterminación y una mecanización de las relaciones del sujeto cognoscente frente a aquello que pretende conocer. De ahí los peligros que se derivan de sobrevalorar el método en el proceso del conocimiento. En efecto, cuando le concedemos un valor en sí, más allá de su función instrumental, el

método se dogmatiza, se convierte para quien lo utiliza en una receta infalible para obtener la verdad y, con ello, deviene la fuente primaria del error.

Hegel consideraba que una descripción similar reproduciría sin duda lo que en realidad había sucedido con todos los métodos de conocimiento anteriores. A su entender, tan sólo en el método dialéctico podía entreverse la posibilidad de escapar a ese destino. Porque, tal como él lo entiende, éste método refleja el dinamismo de su objeto y, al igual que el, no puede renunciar a ese dinamismo sin hacerlo al mismo tiempo a captar la vida de aquello que es objeto del conocimiento. En el sentido pleno del término, existir es prolongar la materialidad más allá de sí misma, gracias al germen de espiritualidad que siempre hay en ella.

El despliegue de lo real no se completa hasta la integración de lo particular en lo general. Como afirma el propio pensador en su *Fenomenología del Espíritu*, “lo verdadero es el todo. Pero el todo es solamente la esencia que se completa mediante su desarrollo. De lo absoluto hay que decir que es esencialmente *resultado*, que sólo al *final* es lo que es en verdad, y en ello precisamente estriba su naturaleza, que es la de ser real, sujeto o devenir de sí mismo” (Hegel, 1806: 16).

Además de en su planteamiento metodológico, en otro aspecto crucial se opone Hegel al kantismo. Se trata de su visión de la conciencia cuyo conocimiento, en lugar de quedar limitado a una intuición basada en la temporalidad, ha de entrar, a su entender, en el ámbito de lo conceptual. En su opinión, “la conciencia es para sí misma su *concepto* y, con ello, de un modo inmediato, el ir más allá de lo limitado y, consiguientemente, más allá de sí misma, puesto que lo limitado le pertenece; con lo singular, se pone en la conciencia, al mismo tiempo, el más allá, aunque sólo sea, como en la intuición espacial, *al lado* de lo limitado” (Hegel, 1806: 55-6).

El surco que marca la diferencia entre ambos pensadores no podía quedar más claro, por otra parte, que con esta apuesta por la intuición espacial, en lugar de hacerlo por la temporal que, como ya se ha reseñado era la clave del planteamiento kantiano. Quizá detrás de la apuesta hegeliana podamos encontrar el objetivo de subrayar la

realidad de la conciencia, a partir de su especialidad y, aún podríamos decir, de su materialidad.

Por lo demás, se produce en Hegel una evidente identificación entre subjetividad y conciencia. Si bien, su concepción de la subjetividad es más compleja de lo usual. Para él, la distancia entre la mera sustancialidad y la aparición del sujeto como tal ha de ir cubriéndose en sucesivas etapas, que engloban cada vez más aspectos supraindividuales. Por eso, cuando se refiere a la primera de esas etapas, nos dice que “la conciencia es *yo* y nada más, un puro *éste*; el singular sabe un puro esto o *lo singular*” (Hegel, 1806: 63).

Momento singular en el devenir de estas etapas es el que supone la aparición de lo se ha llamado la conciencia de la conciencia, la autoconciencia. Unos años después de la publicación de la *Fenomenología del Espíritu*, Wilhem von Humboldt señaló, a este respecto, que la autoconciencia no puede obtenerse sin la intermediación del lenguaje. Para este pensador, era posible la conciencia sin lenguaje, pero la autoconciencia exigía la presencia del mismo, pues la autoconciencia se deriva de una actividad reflexiva sobre los contenidos de la conciencia y ésta última exige el recurso instrumental del lenguaje. Es muy revelador, en este sentido, que Hegel nos diga que para alcanzar la autoconciencia hay que pasar por la conciencia de lo otro, por la conciencia de la distancia que separa al sujeto del objeto o al sujeto de otros sujetos. No se ha cubierto aún la distancia que media entre el yo y el nosotros, pero la conciencia se arma para ese proceso replegándose antes sobre sí misma. Así pues, la idea de la base lingüística de la autoconciencia está presente ya en la *Fenomenología del Espíritu*, de otra manera carecería de soporte real el decurso que Hegel propone para explicar el devenir de la conciencia.

“La conciencia de un otro, de un objeto en general, es, ciertamente, ella misma, necesariamente *autoconciencia*, ser reflejado en sí, conciencia de sí misma en su ser otro. El *proceso necesario* de las figuras anteriores de la conciencia, para la que lo verdadero era una cosa, un otro que ella misma, expresa cabalmente que no sólo la conciencia de la cosa sólo es posible para

una conciencia de sí, sino, además, que solamente ésta es la verdad de aquellas figuras. Pero esta verdad sólo se da para nosotros, aún no para la conciencia. La conciencia de sí sólo ha devenido *para sí*, pero aún no como *unidad* con la conciencia en general” (Hegel, 1806: 103).

Por medio de la actividad intelectual, de la conciencia del objeto puede llegar a derivar el sujeto cognoscente la conciencia de sí. Se trata de un proceso a través del cual el objeto, que en principio había sido percibido como lo opuesto a la conciencia, acaba siendo contemplado como algo con lo que la propia conciencia comparte cierto sustrato. El elemento racional se revela a través del lenguaje, en el caso del sujeto, y de una racionalidad interna que nos indica que no es el puro hacer ni el sinsentido lo que rige la naturaleza y hace que el objeto sea inteligible. De esta forma, el sujeto y el objeto llegarán a ser pensados, no como dos polos opuestos, sino como dos momentos en la proyección de un mismo sustrato racional subyacente. Por eso el sujeto podrá contemplarse a sí mismo en tanto que sujeto de conocimiento, no tanto por lo que proyecta de sí en el objeto para hacer viable su conocimiento, sino por lo que de sí encuentra en el objeto, porque estaba ya presente en el mismo. La conciencia se encuentra consigo misma en el plano de la mediación que, a través de los sistemas simbólicos, ha de establecerse entre el objeto y el sujeto. La autoconciencia se obtiene, en suma, a partir de la conciencia de la conciencia. Hegel considera que “la conciencia tiene ahora, como autoconciencia, un doble objeto: uno, el objeto inmediato de la certeza sensible y de la percepción, pero que se halla señalado para ella con el carácter de lo negativo, y el segundo, precisamente ella misma que es la verdadera esencia y que de momento sólo está presente en la contraposición del primero. La autoconciencia se presenta aquí como el movimiento en que esta contraposición se ha superado y en que deviene la igualdad de sí misma consigo misma” (Hegel, 1806: 108). De donde deriva la idea de una relación dialéctica entre conciencia y autoconciencia. Como resultado de esa dialéctica, la autoconciencia acaba siendo, no sólo el fundamento último de una conciencia que ha superado su propia inmediatez, sino también de la conciencia del objeto. Aquello que hemos llegado a conocer, remontándonos a partir de lo concreto, aparece ahora como fundamento y origen. Lo abstracto se transforma en lo más concreto y lo inmediato manifiesta su naturaleza de mera abstracción al margen de la

actividad mediadora de la autoconciencia. Así, si en el plano de la certeza sensible, la conciencia tomaba como base las referencias sensoriales, en el nivel siguiente, que corresponde a la percepción, la conciencia es conciencia del objeto. De la misma forma que, en el plano del entendimiento lo es de *la fuerza* y en el de la autoconciencia es *la vida* el objeto de la conciencia. Desde esta última perspectiva, la autoconciencia puede permitir al sujeto descifrar el enigma, al revelarle que la conciencia no se pierde en la exterioridad en su intento de conocer los objetos, sino que se adentra en sí misma a medida que avanza en el conocimiento. Por eso afirma Hegel que “el yo es el contenido de la relación y la relación misma; es él mismo contra otro y sobrepasa al mismo tiempo este otro, que para él es también sólo el mismo” (Hegel, 1806: 107).

Queda claro, más aún si cabe después este breve apunte sobre la cuestión, que la interpretación de la obra de Hegel presenta notables dificultades. Pero no lo es menos que, en ocasiones, rindiéndonos de antemano ante ellas, dejamos escapar la posibilidad de compartir con él ideas e intuiciones, a veces geniales, que pueden sernos hoy de gran utilidad a la hora de comprender las diferentes perspectivas desde las que puede abordarse el problema de la conciencia.

Aproximaciones contemporáneas

No es tarea fácil encontrar un hilo conductor que nos ayude a movernos con seguridad entre la diversidad de planteamientos actuales. Las ideas de Hilary Putnam, a propósito de la conciencia y de las relaciones entre la inteligencia artificial y la humana, se cuentan entre las que han obtenido mayor resonancia en esos campos a lo largo de los últimos años. Bien podían, pues, servirnos como tramo inicial al menos de ese hilo conductor que andamos buscando. Una primera aproximación ha de llevarnos a valorar las virtudes e insuficiencias del paradigma funcionalista, cuando se aplica a este campo.

En el enfoque contemporáneo del estudio de la conciencia ha merecido una atención muy especial el estudio de sus estados. A ese respecto, se ha considerado que “la misma propiedad psicológica (por ejemplo, estar iracundo) puede ser una propiedad perteneciente a miembros de miles de especies diferentes, que pueden poseer una

composición física o química completamente distinta (algunas especies podrían ser extraterrestres; y quizá los robots muestren ira algún día). La sugerencia del funcionalista es que la teoría *monista* más plausible que se puede defender (...), la teoría que evita tratar a la mente y a la materia como dos tipos separados de sustancias o como dos reinos separados de propiedades, es la que identifica propiedades psicológicas y propiedades funcionales” (Putnam: 1981, 86).

Se trata, como él mismo nos dice, de evitar el riesgo de caer en una posición dualista, que sustancialice en base a la diferenciación de esas propiedades la existencia de una entidad diferenciada en relación a aquello que es su origen o soporte material. La conciencia misma no sería sino el resultado de un conjunto de propiedades funcionales. De esta forma, si no existe diferencia entre las propiedades psicológicas de una especie a otra, o incluso entre las de un ser vivo y las de un ordenador, tampoco debería existir diferencia alguna por lo que se refiere a la conciencia. En todo caso, de existir alguna diferencia, ésta sería de intensidad o grado, pero no de naturaleza. Podría haber estados psicológicos inviables, dada la configuración de una máquina o un ser, pero, en aquellos que esas entidades pudieran compartir, no existiría diferencia alguna.

No obstante, Putnam nos advierte de inmediato acerca de cierto tipo de dificultades inherentes a esta concepción. A su entender, “cuando se piensa en estados psicológicos puros y relativamente abstractos, por ejemplo, en lo que he denominado *creencias entre paréntesis*, esto es, un pensamiento considerado en su contenido *nocional*, o en estados emocionales tan difusos como estar celoso o iracundo, su identificación con estados funcionales de todo el sistema parece suficientemente plausible; pero cuando se piensa en experimentar una cualidad dada, por ejemplo, experimentar una tonalidad particular de azul, la identificación ya no es plausible” (Putnam: 1981, 87). Es decir, que cuando nos aproximamos a cierto grado de diferenciación o sutileza, la citada teoría deja de ser satisfactoria. Habría que añadir, a este respecto, que tales estados, así como otros muchos de mayor complejidad aún, son los que de una manera más específica podemos identificar con lo propio del ser humano. A través de ello se pone en evidencia que merecería ser cuestionada toda hipótesis interpretativa que no nos ofreciese una explicación suficiente de los mismos.

Con todo, la correlación entre estados funcionales y contenidos de conciencia es considerada comúnmente como algo obvio. En efecto, parece evidente que no podrían existir contenidos de la conciencia sin un soporte físico para los procesos que los ocasionan. Estos procesos pueden describirse como la génesis de estados funcionales. Como decimos, pese a que todo esto parece rozar la obviedad, Putnam se opone a la simplificación de un problema que, dada la interpretación que hacemos del mismo, pierde por completo su perfil como tal. Para expresarlo con sus propias palabras, “la correlación no es una cuestión muy debatida, pues todo el mundo sabe que hay al menos una correlación. La identidad sí que se discute, porque es ahí donde radica el problema”. Sin entrar a sopesar la dificultad o facilidad de probar la tesis de la identidad, se centra en demostrar que “hasta la *correlación es problemática, y no en el sentido de que exista evidencia a favor de la no-correlación, sino en el sentido epistemológico de que aunque exista una correlación, nunca podremos averiguar cuál es ésta*”. En conclusión, como añade el propio autor, “el problema no radicará en asumir el materialismo, sino en el hecho de que creamos que hay al menos una correlación” (Putnam: 1981, 89).

A la hora de establecer esa correlación con frecuencia se olvida, según Putnam, la dificultad inherente al intento de relacionar procesos, cada uno de ellos con un dinamismo propio, a los que se pretende reducir a un patrón uniforme. A su juicio, “el problema consiste en que la psicología divide los eventos mentales de una forma excesivamente discreta. He aquí la sensación de azul: en este momento empieza, en este otro acaba. Pero las redes causales no son discretas. No hay un único evento físico que sea el correlato de la sensación.

Si la teoría de la identidad está en lo cierto, el estado sensorial Q es idéntico a algún estado cerebral. Un realista metafísico no puede considerar esta identidad como un asunto de convención o de decisión o como si se tuviese un componente convencional idéntico al estado cerebral Q. La opinión es que, como *cuestión de hecho*, vivimos en un mundo en el cual lo que experimentamos como caracteres cualitativos de las sensaciones son en realidad *las mismas propiedades* que algunas de las propiedades que nos encontramos en otros campos como propiedades físicas de eventos cerebrales

(O, mejor dicho, en los cuales la propiedad de experimentar una sensación con cierto carácter cualitativo es exacta y realmente la propiedad de hallarse en cierto estado cerebral)” (Putnam: 1981, 93-4).

Sin embargo, no basta con postular la identidad entre estados de conciencia y estados funcionales. Se trata de una hipótesis deudora de ciertos presupuestos ontológicos y cuya veracidad ha de ser comprobada. Aunque para él los problemas provienen ante todo de su relación con el realismo metafísico, que hacía de las entidades abstractas universales el único soporte válido para el conocimiento. El rango concedido a tales universales los situaba por encima de los contenidos derivados de la experiencia, pese a que sin estos últimos resultara imposible, como señalaba Kant, la elaboración de los conceptos. Por ello sostiene Putnam que “no es la teoría de la identidad la que de por sí nos conduce a dificultades, sino la teoría de la identidad considerada en conjunción con el realismo metafísico –es decir, considerada en conjunción con lo que he denominado perspectiva ‘externalista’ con respecto a la naturaleza de la verdad”. Frente a esa concepción existen diversas alternativas; él mismo nos recuerda que “se puede eludir el compromiso con esa perspectiva” y cita como ejemplo a Carnap, quien habría defendido “que el discurso acerca de objetos físicos es en realidad discurso acerca de sensaciones, si bien muy derivado, y que la decisión de afirmar que un particular estado cerebral es idéntico a un estado sensorial Q es, en realidad, la decisión de modificar de cierta forma el lenguaje del discurso acerca de propiedades físicas, de cambiar nuestro concepto de la propiedad física en cuestión” (Putnam: 1981, 95).

En efecto, al referirnos a los objetos físicos lo hacemos a través de la imagen que de los mismos tenemos presente en nuestra conciencia. Esa *imagen* corresponde a los contenidos perceptuales asociados al objeto. Se puede dar un paso más allá y considerar que los estados mentales, asociados a la percepción o a cualquier otro proceso cognitivo, son en realidad el soporte último de los *universales*. Esos modelos básicos, cuyo papel sería esencial en los procesos cognitivos al permitirnos agrupar bajo un mismo entramado conceptual contenidos heterogéneos, tendrían así una base fisiológica y no meramente cultural.

La existencia de esas estructuras objetivas es incuestionable para quienes defienden una posición realista, en el sentido tradicional del término. Para estos últimos, los universales son a la vez la base del conocimiento y uno de los fundamentos del orden que percibimos en la realidad. Son, por tanto, más reales que los mismos objetos percibidos. Los estados mentales específicos de cada individuo, a los que llamamos *qualia*, responderían a esas estructuras y constituirían su máximo exponente. Precisamente, por esa misma razón Russell pensaba que los *qualia* son los universales en sí. Ciertamente, cuando distintos objetos remiten a un mismo estado de conciencia, podemos inferir que todos ellos se relacionan con el mismo universal.

A juicio de Putnam, sería imposible distinguir los estados mentales de distintos entes, por más que la configuración funcional que diera lugar a los mismos fuera específica de cada uno de ellos. De esta forma, “si efectivamente hubiera robots que fueran funcionalmente isomórficos a nosotros, y trabajásemos, discutiésemos e incluso entabláramos amistad con algunos de ellos, no dudaríamos ni por un momento que fuesen conscientes (Aún así, podríamos quedar perplejos ante la cuestión de si sus *qualia* son como los nuestros; pero esta cuestión no se plantearía más a menudo que la cuestión de si los perros o los murciélagos tienen los mismos *qualia* que nosotros)” (Putnam: 1981, 107). Cualquiera de esos entes dispondría, por tanto, de *universales* correlativos a sus respectivos *qualia*, y no podríamos distinguirlos de los de cualquier otro ente capaz de estar en un mismo estado funcional. Con ello, el soporte último de la objetividad y universalidad del conocimiento tampoco sería patrimonio específico de nuestra especie, sino que permanecería compartido por todo ente susceptible de tener *qualia* análogos a los nuestros.

Por otra parte, como vemos, Putnam subraya la familiaridad como propiedad psicológica que nos lleva a reconocer o no a un ente como semejante a nosotros. El antropomorfismo sigue estando presente en nuestra interpretación del pensamiento o en la atribución de la conciencia. “Nuestro mundo es un mundo humano, y la respuesta a qué cosas son conscientes o no, o a qué cosas experimentan sensaciones o no, o a qué cosas son cualitativamente semejantes o no, depende en última instancia de nuestros juicios humanos con respecto a la semejanza y a la diferencia” (Putnam: 1981, 108).

Pese a todo, habría que tener presente que, cuando pensamos que cualquier ser capaz de percibir algo que un cerebro humano está percibiendo, estaría en el mismo estado funcional que ese cerebro, pueden planteárenos varias cuestiones. Algunas de ellas, es bien cierto, son de orden emocional y tienen que ver con la imagen que el ser humano tiene de sí mismo. En efecto, siempre habrá quien no admita o se sienta incluso agredido por la simple idea de otro ser o de una máquina situados en un estado de conciencia indiferenciable del de un humano. Otras dificultades, sin embargo, trascienden los prejuicios culturales para entrar en un orden epistémico. Se refieren al conocimiento de los procesos cognitivos humanos, a su funcionamiento complejo, continuo y holístico, en lugar de un supuesto modo de proceder discreto. Porque no percibimos cualidades simples y aisladas. Si algo ha quedado claro después de los avatares de la teoría gestaltista es que la percepción humana es un proceso intencional que implica a distintas estructuras, desde la conciencia a la memoria, pasando por las relacionadas con la conducta emocional. Así, no puede ser idéntico un estado de conciencia determinado de un ser humano al de, digamos, un murciélago porque las estructuras cerebrales y los respectivos contenidos de ambos cerebros no son equipolentes. No es esta, en cualquier caso, una conclusión que venga dictada por ninguna suerte de orgullo antropocéntrico. La cisura entre ambos cerebros marca una distancia insalvable en una y otra dirección. Esa cisura es menos pronunciada cuando nos referimos a otros animales, como los gorilas o los chimpancés, pero también existe. No sólo por la diferencia en el tamaño y configuración del neocórtex, sino también por el uso que, desde su aparición en el proceso evolutivo, damos a tales estructuras.

En *Representación y realidad*, el filósofo norteamericano realiza un encomiable esfuerzo autocrítico, distanciándose de manera definitiva de sus primeras posiciones en relación al funcionalismo. Como él mismo nos dice, es posible que sea “el primer filósofo que propuso la tesis de que el ordenador es el modelo apropiado para la mente”. A continuación, señala cómo estableció la denominación de *funcionalismo* a su “versión de esta teoría” y cómo ésta “se convirtió en la concepción dominante –o en la ortodoxia, como dicen algunos- dentro de la filosofía contemporánea de la mente”.

Ahora se muestra convencido de “que la analogía de la mente con el ordenador, llámese ‘teoría computacional de la mente’, ‘funcionalismo’, o como se quiera, no da respuestas a la pregunta que muchos filósofos (y muchos científicos cognitivos) deseamos responder: ‘¿cuál es la naturaleza de los estados mentales?’” (Putnam, 1988: 13-4).

Desarrollando paso a paso las bases de su nueva interpretación postula, en primer término, que “los estados mentales no pueden ser ‘programas’, pues dos sistemas físicamente posibles pueden hallarse en el mismo estado mental y tener, no obstante, ‘programas’ diferentes” (Putnam, 1988: 18). En consecuencia, esa interpretación, usual dentro de los círculos funcionalistas, ha de ser descartada. De este modo, la conciencia habría de ser algo más que un mero epifenómeno asociado al funcionamiento de un programa. No es viable, por tanto, reducirla a una interpretación algorítmica.

Recurriendo a un concepto hoy en boga, Putnam se remite para proseguir la discusión con el funcionalismo a las concepciones de Chomsky y Fodor.

A su entender “no hay esperanzas de hallar una teoría sobre la identidad del significado o la referencia” que pueda aplicarse a partir de cierto nivel de complejidad “y que sea independiente de nuestra explicación de la ‘inteligencia general’” (Putnam, 1988: 39). Su alusión se refiere, de una manera particular, a los excesos a que ha conducido el desarrollo de las ideas de Fodor.

Aristóteles, en su *De interpretatione*, estableció que cuando comprendemos una palabra, la asociamos a un concepto. En esa misma línea se han movido muchos autores posteriores, como John Stuart Mill, Bertrand Russell, Frege, Carnap, etc., todos ellos hicieron de una u otra forma equivaler los conceptos a las representaciones mentales.

El gran problema siempre ha sido, a ese respecto, justificar de modo suficiente la relación entre símbolos, representaciones mentales y objetos del mundo exterior. Como el mismo Putnam nos recuerda, existe una notable incertidumbre en este campo pues, “lo que hace creíble que la mente (o el cerebro) piense (o compute) por medio de

representaciones es que todo el pensamiento que conocemos se vale de representaciones. Pero ninguno de los métodos de representación que conocemos – discurso oral, escritura, pintura, grabado en piedra, etc.- tiene la propiedad mágica de no admitir representaciones diferentes del mismo significado” (Putnam, 1988: 48-9). No puede probarse, antes al contrario, ninguna ligazón natural entre el significante y el significado. De tal forma que el pensamiento se vale de recursos que de suyo son susceptibles de ser intercambiados o sustituidos por otros, ya que nada en ellos remite de forma unidireccional e inequívoca al significado que les atribuimos.

La existencia de un lenguaje interno de la mente, al que no sin ironía se ha denominado *mentales*, no impediría que el significado o la referencia siguiesen envueltos en la equivocidad. El lenguaje y el pensamiento encontrarían en ese instrumento intermedio la forma de articularse pero, como ya sucedió con hipótesis análogas a lo largo de toda la historia de la filosofía, el *mentales* tan sólo supondría un desplazamiento del problema y no su solución, además de inducirnos, si aceptamos su existencia, a correr el riesgo de un *regresum ad infinitum*. Pues, la *lingua mentis* exigiría la presencia de otra instancia intermedia entre ella y las redes neuronales correspondientes y a su vez, esa nueva instancia podría exigir otra instancia intermedia para explicar el nivel de complejidad de la comunicación, y así sucesivamente.

A ello se añade que, con *mentales* o sin él, no parece haber una forma clara de dejar bien asentada la ligazón entre lenguaje y mundo de la que, desde Aristóteles, tanto se ha hablado. En cierto modo, esa es la lección más importante que podemos extraer de la lectura del diálogo *Crátilo*, que su maestro Platón dedicó a discutir el origen del lenguaje y su utilidad como instrumento del conocimiento. Pero Aristóteles, quizá pensando en un modo de huir del peligro de misticismo subyacente en el planteamiento platónico que hace de los universales, a un mismo tiempo, objeto de intuición y único fundamento objetivo del ser y el conocer, se embarcó en la estrategia de justificar a través de los universales un modo de articular el valor objetivo de los modos de significar.

Para justificar su oposición a ese enfoque, Putnam se apoya en el hecho de la “división lingüística del trabajo”. Para él, “el lenguaje es una forma de actividad cooperativa y no una actividad esencialmente individualista. La imagen aristotélica es errónea cuando supone que todo lo necesario para el uso de la lengua está almacenado en la mente individual”. Como es sabido, no es eso exactamente lo que sostenía Aristóteles. Para éste, el *entendimiento agente*, común a la especie, es el responsable de la base universal para el pensamiento y el lenguaje. No obstante, la existencia de tal entidad es postulada más que demostrada por el filósofo griego. Frente a su hipotética existencia, lo constatable en todo caso es el generalizado dominio del lenguaje entre los humanos. El funcionamiento de éste no podría explicarse, en términos aristotélicos, sin la competencia individual. En cualquier caso, Putnam recalca que “ninguna lengua funciona, en realidad, de esa manera” (Putnam, 1988: 54).

Parece incongruente con nuestro conocimiento actual del modo en que funcionan los procesos cognitivos, imaginar que todos los términos del lenguaje hayan de estar asociados a contenidos conceptuales precisos. Por otro lado, como señala Putnam, la tesis del prototipo perceptual también puede presentar dificultades, cuando la tratamos de aplicar en contextos como el de la traducción de una lengua a otra. “Los prototipos perceptuales, si bien son importantes para la psicología, no son *significados*, ni siquiera significados ‘estrictos’” (Putnam, 1988: 81-2). El prototipo perceptual no basta para establecer el significado.

Así pues, si no podemos identificar prototipo perceptual y significado, hemos de orientarnos hacia la tesis que concede primacía a la dimensión social en el establecimiento de la relación entre el significante y el significado. No tiene sentido conceder prioridad a las representaciones mentales cuando, como hemos visto antes, los prototipos perceptuales tienen un contenido que varía en función de la dimensión social del significado.

En resumen, insiste el filósofo norteamericano, “no podemos individualizar los conceptos y las creencias sin hacer referencia al *entorno*. Los significados no están ‘en la cabeza’. El resultado de nuestra discusión es que las actitudes proposicionales, como

las llaman los filósofos –es decir, cosas tales como *creer que la nieve es blanca* y *estar seguro de que el gato está sobre el felpudo*- no son ‘estados’ del cerebro humano y del sistema nervioso aislados del entorno social y no humano. *A fortiori*, no son ‘estados funcionales’, o sea estados definibles en términos de parámetros que forman parte de una descripción del *software* del organismo”. Para Putnam, al menos en este sentido, la propuesta funcionalista es incoherente. En efecto, su conclusión es que, “*el funcionalismo, entendido como la tesis de que las actitudes proposicionales son meros estados computacionales del cerebro, no puede ser correcto*” (Putnam, 1988: 119-20).

Por eso él propone una orientación diferente, con el objetivo de aclarar cuál es en realidad el *modus operandi* de la conciencia. En su planteamiento puede apreciarse con claridad la profundidad del giro que sus ideas han experimentado. Nos dice así que, lo que suele suceder es que, “en vez de pensar que las actitudes proposicionales tienen una realidad fenomenológica que surge de la posibilidad de preguntarnos si comprendimos correctamente un texto o a una persona, buscamos una reducción de las actitudes proposicionales a algo más ‘básico’ en nuestro sistema de metafísica científica. Buscamos algo definible en términos no intencionales, algo a partir de lo que podamos construir un modelo, algo que *explique* la intencionalidad. Y esto –el ‘proceso mental’- es precisamente lo que no existe” (Putnam, 1988: 121).

Como ya hemos señalado, la prioridad que puede concederse en el establecimiento del significado a los estereotipos por encima de las actitudes proposicionales, sitúa el aspecto social y cultural como espacio predominante. La comprensión de un texto o del discurso de un hablante es mucho más que la aceptación de que tales mensajes se expresan con corrección formal, empleando unos símbolos y unas reglas determinadas. Por tanto, como decíamos antes, una interpretación algorítmica del funcionamiento de la conciencia está fuera de lugar, ya que no podría ofrecernos una explicación suficiente de cómo se produce la atribución del significado a los elementos que utilizamos en nuestro lenguaje. En definitiva, “el hecho de que no podamos interpretar un discurso a menos que seamos capaces de continuarlo indica que un algoritmo que interpretara un discurso arbitrario tendría que ser lo suficientemente inteligente para examinar todos los posibles discursos racionales, semirracionales y

poco racionales pero aún inteligibles que pudiesen construir de modo físicamente posible todas las criaturas físicamente posibles. ¿Es probable que exista un algoritmo semejante?” (Putnam, 1988: 140). No parece que la respuesta pueda ser hoy por hoy afirmativa.

Por otro lado, continúa Putnam, “si se recuerda que el único ‘asidero’ que poseemos de la noción de ‘idéntica creencia’ es, en realidad, la práctica de la interpretación, no hay ninguna razón entonces para creer que existe un *estado computacional* en el que se hallan todos los seres humanos posibles que piensan que ‘hay muchos gatos en el vecindario’. Y si no existe un *estado computacional* único en el cual se encuentren todos, tampoco es posible que se encuentren en un estado *neuroquímico* relevante” (Putnam, 1988: 163). Así pues, si el significado que atribuimos a esos u otros términos no está predeterminado por unas estructuras fijas, con una base neuroquímica y comunes a toda la especie, no es coherente pensar que exista una correspondencia entre los contenidos de la conciencia de cada sujeto ni que todos ellos puedan encontrarse en un mismo estado computacional. La vinculación del significado de las palabras a su uso, le permite alejarse de todo presupuesto metafísico en torno al establecimiento del significado. A ello se añade que el conocimiento del proceso nos revela un contenido que va más allá de nuestras expectativas iniciales. “En suma, el intento de examinar el ‘significado’ o la referencia fracasa por la misma razón que fracasa el intento de examinar la razón misma: la razón puede trascender aquello que examina” (Putnam, 1988: 182). Siguiendo una pauta análoga a la que estableciera Wittgenstein en sus *Investigaciones filosóficas*, nos emplaza ahora a abandonar los intentos de explicación neuroquímica, para adentrarnos en el camino de la indagación acerca del uso del lenguaje.

En todo caso, cabría preguntar si trasciende la conciencia aquello que examina. Por lo que llevamos visto hasta ahora, tendríamos que concluir que así es. Y, si como parece es eso lo que sucede, sería posible preguntarnos, evocando una de las líneas de aproximación sugeridas por Hegel, si también se trasciende a sí misma cuando se toma como objeto de estudio.

Habría entonces, de la misma manera que se ha intentado definir el significado de las palabras a partir de su uso, que tratar de definir la conciencia por medio del análisis de su cometido. ¿Qué hacemos mediante la conciencia? ¿Cómo hacemos aquello que hacemos con ella? Estas siguen siendo las grandes preguntas, a las que desde diferentes líneas de estudio nos aproximamos pero que aún no hemos sido capaces de responder desde ninguna de ellas.

En todo caso, vemos cómo va asentándose con firmeza la conclusión de que “debe haber un ingrediente esencialmente *no-algorítmico* en la actuación de la conciencia”. (Penrose: 1989, 505). Esa y no otra es la razón última que avala la convicción de la imposibilidad de reproducir de forma mecánica el funcionamiento de la conciencia. Desde esta perspectiva, ninguna máquina podrá nunca llegar a pensar en la misma forma en que puede pensar un cerebro humano.

De forma análoga, tampoco podemos inferir los contenidos de la conciencia guiándonos por sus manifestaciones externas. Un ejemplo de ello, tan explícito como dramático, es el que Penrose nos ofrece al evocar la utilización del *curare* en cirugía. La pauta general de inferir los contenidos mentales a partir de la conducta del sujeto, se muestra inadecuada si pretendemos hacer de ella un canon explicativo de aplicación universal. Si extraemos de esto todas las consecuencias, tampoco la imitación de la conducta humana en un robot debería tomarse como prueba de un comportamiento propiamente humano. Así, la búsqueda del elemento clave del comportamiento humano nos remite de nuevo a la peculiaridad de la conciencia.

No nos sentimos aislados en el interior de nuestra mente porque compartimos la vida y sus cuidados con otros seres a los que reconocemos análogas inquietudes a las nuestras. ¿De qué forma alcanzamos esa convicción? En contra de lo que sugería Putnam, al decirnos que es tan sólo la familiaridad lo que nos proporciona esa certeza, es la comunicación simbólica, en particular la lingüística, la que nos permite disponer de los medios necesarios para alcanzarla. A través de nuestro uso del lenguaje estamos pasando de forma cotidiana el *Test de Turing*.

Pero no vayamos más allá sin reconocer que, como puede intuirse por la alusión precedente, pese a estar expuesta sobre el fondo de un planteamiento dualista, las tesis que plantea Penrose son, cuando menos, sugerentes. Por lo demás, no parece ser de recibo en la actualidad entender que la mente sea algo con sustancialidad propia e independiente de la que atribuimos al cerebro. Aunque tampoco sea necesaria tal presuposición dualista para comprender y compartir el trasfondo de la interpretación de Penrose. En todo caso, quizá el límite haya de situarse ahora, no tanto entre un monismo y un dualismo matizados, cuanto entre un materialismo ramplón que se acoge sin remedio a unas insatisfactorias propuestas mecanicistas y un monismo que huye del simplismo y nos sitúa en el camino de una interpretación satisfactoria de la conciencia. Sea como fuere, según nos ha enseñado Gadamer, “la dialéctica del límite es que sólo es en cuanto que se supera” (Gadamer, 1975: 417). En suma, tan sólo en la medida en que sobrepasemos las declaraciones de principio y nos adentremos en un terreno en el que sea por fin factible una interpretación coherente de la conciencia, podremos tener cierta confianza en haber trascendido esa limitación empobrecedora.

Por lo demás, podría considerarse que en la actualidad el estudio de la conciencia es el límite focal de toda una serie de disciplinas. Representa una frontera, aunque quizá no acabe de ser percibida como tal hasta que no se alcance una explicación suficiente de la misma. Entonces tendremos plena certeza de su posición limítrofe así como de haber trascendido dicho límite.

En este momento, cuando nos dirigimos ya al final de estas páginas, puede ser clarificador hacer una breve recapitulación. Para ello tal vez sea de utilidad evocar el contenido de la obra *El misterio de la conciencia*, de John Searle. Se trata de un ensayo de gran interés, tanto por la síntesis que nos ofrece de los planteamientos actuales cuanto por el análisis crítico de los mismos que en sus páginas se realiza. Podríamos, pues, seguir su línea expositiva para recordar las ideas que hasta ahora hemos debatido y avanzar algunas propuestas a modo de conclusión.

Una de las cuestiones sobre las que resulta inevitable volver es la que se interroga por las similitudes y diferencias que pueden existir entre el cerebro y el computador. En efecto, como hemos visto es un planteamiento recurrente el de si puede

o no pensar una máquina. A propósito de él, Searle nos recuerda que “el cerebro es una máquina, una máquina biológica, y puede pensar” (Searle: 1997, 13). En todo caso, esta respuesta deja abierta una vez más la cuestión de si alguna vez se podrá construir una máquina capaz de pensar de forma que su pensamiento sea indiscernible del característico de un cerebro humano. Si bien, la respuesta que pueda darse a esta pregunta dependerá en buena medida de qué entendamos por pensar. Si se trata de computar o realizar operaciones, es evidente que tanto ciertas máquinas creadas por nosotros como el propio cerebro humano, son capaces de *pensar*. Pero, si señalamos a la conciencia como base y rasgo diferencial del pensamiento, la distancia entre el cerebro humano y una máquina sigue siendo muy grande. En este sentido, en contra de lo sostenido por la versión fuerte de la Inteligencia Artificial, Searle se une a otros pensadores que hemos mencionado aquí, al sostener que la conciencia no puede ser nunca un mero efecto derivado del funcionamiento de un programa de ordenador (Searle: 1997, 14). Para entender la base de su planteamiento hay que distinguir entre el manejo de símbolos según ciertas reglas definidas de antemano y la comprensión del significado de esos símbolos, más allá de las reglas que definen la correcta combinatoria de los mismos. Esta distinción, que algunos juzgarán problemática, no deja por ello de llevarnos a aceptar que los seres humanos hacemos un uso de los símbolos que no tendría sentido al margen de la existencia de la conciencia. Para Searle, el cerebro es “un órgano como cualquier otro; se trata de una máquina orgánica. La conciencia es causada por el nivel básico de los procesos neuronales en el cerebro y es en sí misma una característica del cerebro. Puesto que es un rasgo que emerge de ciertas actividades neuronales, podemos pensar que es una ‘propiedad emergente’ del cerebro” (Searle: 1997, 17-8). Sin embargo, como ya se ha dicho, la conciencia no puede ser explicada en términos algorítmicos. No se trata de un programa que un computador pueda ejecutar. Los programas informáticos manejan símbolos con una apabullante rapidez y eficacia, pero la conciencia no puede ser producida por ninguno de ellos. Tampoco se trata, por otra parte, de un epifenómeno asociado al funcionamiento del cerebro. Por el contrario es un rasgo primordial del cerebro humano, aunque sabemos que los animales más próximos a nosotros en la escala evolutiva dan muestras de poseer también conciencia, al menos en cierto grado.

Otro de los asuntos que han suscitado mayor atención ha sido el llamado *problema de los qualia*. Como ya hemos expuesto, esta cuestión alude a la dificultad de conocer los estados de conciencia de otra persona. “¿Cómo es posible mediante una descripción física, objetiva y cuantitativa, describir de qué manera las neuronas causan las cualitativas y privadas experiencias subjetivas?” (Searle: 1997, 28). Searle sitúa aquí el núcleo del tradicional problema de las relaciones mente-cuerpo. Esta perenne cuestión ha de interpretarse, a su juicio, a través de la búsqueda de una solución al problema de los *qualia*, porque de lo que se trata ahora es de explicar cómo órganos, no idénticos pero de análoga complejidad, producen estados de conciencia diferentes. Así, aunque el soporte físico sea análogo, los contenidos mentales tienen una especificidad propia en cada sujeto. De la misma forma, pese a que los símbolos que empleamos para la construcción, representación y transmisión del pensamiento son los mismos, los contenidos mentales que hacemos corresponder a esos símbolos se relacionan con experiencias y contenidos que son específicos del individuo. Pero no hay detrás de estas palabras una defensa del solipsismo. Antes al contrario, la posibilidad de la comunicación se basa en la analogía y la diferencia subyacente a la misma, más que en la identidad. Cerebros idénticos, con idénticos contenidos de conciencia, no tendrían necesidad alguna de comunicarse pues nada habría que comunicar.

Pero quizá el mayor de los méritos del libro de Searle que estamos comentando radique en el análisis y discusión que nos ofrece de las más relevantes teorías acerca de la conciencia humana. Nos facilita así, por ejemplo, un detallado resumen de los planteamientos de Edelman, centrándose en las nociones clave manejadas por este investigador: a) la noción de mapas neuronales, b) la teoría de la selección neuronal, que viene a ser una traslación de la hipótesis básica de la teoría darwiniana aplicada al proceso de selección en el que unos grupos de neuronas acaban imponiéndose a otras, y c) la idea de reentrada, descrita como un proceso que permite a diversos mapas neuronales trabajar en paralelo. Por tanto, la *reentrada* no ha de identificarse con un simple *feed-back*. El mecanismo de reentrada hace posible la construcción de un mapa global, que permite al individuo, no sólo disponer de categorías perceptivas, sino también coordinar los procesos perceptivos con sus acciones.

Basándose en estos presupuestos, Edelman construye un modelo explicativo en el cual la memoria -entendida como un proceso continuado y no como un resultado-, la concepción del cerebro como un sistema de aprendizaje, las habilidades cerebrales para discriminar el yo del no-yo, los mecanismos de categorización de acontecimientos y formación de conceptos, así como los que se encargan de relacionar los procedimientos perceptivos con las categorizaciones, son los recursos primordiales para explicar la aparición y funcionamiento de la conciencia en los seres humanos (Searle: 1997, 40 y ss.).

En cuanto a su comentario de las ideas expuestas de Roger Penrose, hemos de decir de antemano que tiene un tono mucho más crítico que el que encontramos en las páginas dedicadas a comentar los planteamientos de Edelman que, como no podía ser de otra manera, Searle encuentra muy prometedores. En efecto, ni la demostración del sinsentido de las tesis de la IA, tomando como base una aplicación del Teorema de Gödel, ni tampoco la apelación a la mecánica cuántica que realiza le parecen a Searle partes de la solución del *misterio* de la conciencia. En primer término porque, sin entrar a discutir la tesis de fondo, los argumentos de Penrose son inconsistentes a juicio de Searle. En segundo lugar, porque la explicación, que implica el uso de conceptos extraídos de la física cuántica, le parece aún más misteriosa que aquello que pretende explicar (Searle: 1997, 55 y ss.).

Pero es en otra de las obras recientes de Searle, *Mente, Lenguaje y Sociedad*, donde encontramos algunos de los planteamientos que pueden ser más esclarecedores. Así, en el citado ensayo, el filósofo norteamericano nos ofrece una definición de las características estructurales de la conciencia que tiene un gran interés para ser empleada a modo de compendio esquemático. Para él esas características son las siguientes (Searle:1999, 73 y ss.):

1. Los estados de conciencia sólo existen como experiencias de un sujeto agente.
2. La conciencia se presenta como una experiencia unificadora o integradora.
3. La conciencia posibilita la interacción eficiente con el mundo exterior.

4. Existe una relación entre estados de conciencia y estados de ánimo. La depresión o la euforia, en circunstancias normales, son resultado de una interacción con estados de conciencia.
5. En sus formas no patológicas, los estados de conciencia aparecen como estados estructurados o estructurantes.
6. Hay una relación entre los estados de conciencia y los variados grados de atención que un sujeto puede experimentar.
7. La conciencia tiene unas condiciones límite.
8. Nos permite disponer de distintos grados de familiaridad en relación a las experiencias que podemos tener.
9. Toda experiencia es relacionada en la conciencia con experiencias anteriores. Nunca es tratada como experiencia aislada.
10. Los estados de conciencia aparecen asociados a distintos grados de placer o displacer en el sujeto.

Por añadidura, en las páginas que dedica al estudio de la conciencia en esta obra, Searle insiste en la importancia de esclarecer dos cuestiones fundamentales en relación a ella. Por una parte, el problema del *binding*, que como se sabe consiste en intentar explicar cómo puede la conciencia construir una unidad de experiencia tomando como base una gran cantidad de elementos dispersos (Searle:1999, 80 y ss.). Por otro lado, nos presenta también como cuestión de primordial importancia la relación entre conciencia e intencionalidad. Para él no se trata de una relación accidental sino que, por el contrario, sin considerar que la vinculación entre ambas es esencial no se puede explicar el funcionamiento de la conciencia (Searle:1999, 86 y ss.).

Insiste en que aunque el cerebro humano puede usar algoritmos matemáticos, aunque no consiste en tales algoritmos. De la misma forma que, si bien el ordenador con el que se quiere simular el funcionamiento del cerebro puede manejar algoritmos, ello no significa que éste consista en tales algoritmos (Searle:1999, 76).

Mención aparte merece su exposición crítica de otras teorías que niegan la existencia de la conciencia, en particular la defendida por Daniel Dennett. Este autor utiliza el concepto de *meme*, tomado de Dawkins, para decirnos que lo que llamamos

conciencia no es otra cosa que un conjunto de esas unidades de transmisión o de imitación cultural. De esta manera, el ser humano quedaría reducido a la condición de un autómata que recibe y emite información, en función de parámetros preestablecidos culturalmente (Searle:1999, 104 y ss.).

La argumentación de Searle en contra de estas chocantes ideas, se centra en la imposibilidad de distinguir, siguiendo los planteamientos de Denett, entre apariencia y realidad. Desde esa perspectiva, se necesitaría, en efecto, una tercera persona para aportar coherencia a cualquier interpretación fenoménica. Es más, podríamos añadir que esa tercera persona quedaría investida de una capacidad objetivante que en realidad no podría serle atribuida con fundamento. Con lo cual, el problema de la consistencia de la realidad o el de la objetividad del conocimiento no podrían resolverse sino recurriendo a la vieja hipótesis de Berkeley y llamando en nuestra ayuda a una superconciencia, que el pensador británico identificaba con Dios. Nada más lejos de la perseguida objetividad científica, podríamos concluir.

En sus páginas finales, esta obra nos ofrece un breve compendio de las propuestas del autor en relación al problema de la conciencia. A mi juicio, merecen destacarse dos de ellas. En primer lugar, que para dejar de hablar de la conciencia en términos de misterio o enigma y empezar a hacerlo en términos de problemas científicos que tienen o pueden tener una explicación concreta, hay que asumir que el objetivo final es explicar cómo determinados procesos neurológicos en el cerebro humano, pueden producir estados de conciencia subjetivos, conocimientos o experiencias y de qué manera se materializan esos estados en las estructuras cerebrales (Searle:1999, 192). Por otro lado, reviste particular interés el planteamiento de Searle en torno al dualismo, pues, tal y como se ha indicado en estas mismas páginas, él piensa que es posible rechazarlo sin caer en un reduccionismo. En definitiva, se puede ser monista sin ser reduccionista, sin negar su entidad peculiar a los fenómenos conscientes (Searle:1999, 211 y ss.).

Bibliografía

- CHANGEAUX, J.P. (2002): *L'homme de vérité*. Paris, Odile Jacob.
- GADAMER, H.G. (1975): *Verdad y método*, trad. A. Aparicio y R. De Agapito, Salamanca, Sígueme: 1977.
- HEGEL, G.W.F.(1807): *Fenomenología del Espíritu*, trad. W. Roces, México, FCE: 1985.
- HOUDE, O., B. Mazoyer et Tzourio-Mazoyer, N. (2002) : *Cerveau et psychologie*, Paris, PUF.
- ORTEGA Y GASSET, J.: “Prólogo”, a la edición española de HEGEL, G.W.F.: *Lecciones sobre la Filosofía de la Historia Universal*. Trad. J. Gaos. Madrid, Alianza: 1980.
- PENROSE, R. (1989): *La nueva mente del emperador*. Trad. J. García. Madrid, Mondadori, 1991.
- POPPER, K. (1994): *El cuerpo y la mente*. Trad. O. Domínguez. Barcelona, Paidós: 1997.
- PUTNAM, H. (1981): *Razón, verdad e historia*. Trad. J. M. Esteban. Madrid, Tecnos: 1988.
- PUTNAM, H. (1988): *Representación y realidad*. Trad. G. Ventureira. Barcelona, Gedisa, 1990.
- KANT, I. (1787): *Crítica de la Razón Pura*. Trad. P. Ribas. Madrid, Alfaguara. 1989.
- SEARLE, J. (1997): *The Mystery of consciousness*, London, Granta Books. 1998.
- SEARLE, J. (1999): *Mind, Language and Society*. London, Phoenix.
- STENGER, V.J.(1992): “The Myth of Quantum Consciousness”, *The Humanist*, LIII, 3 (may/june 1992).
- THOM, R. (1977): *Estabilidad estructural y morfogénesis*,.Trad. A.L.Bixio. Barcelona, Gedisa. 1987.
- VIVES, L. (1538): *Tratado del alma*. Trad. J. Ontañón. Madrid, Ediciones de la lectura. 1925.