



**UNIVERSIDAD NACIONAL
AUTÓNOMA DE MÉXICO**

FACULTAD DE FILOSOFIA Y LETRAS

**SUBJETIVO Y OBJETIVO. CONCEPTOS
RELEVANTES PARA LA CONSTRUCCIÓN DE
LA CIENCIA DE LA CONCIENCIA**

T E S I S

QUE PARA OBTENER EL TITULO DE
LICENCIADO EN FILOSOFIA

P R E S E N T A

OSCAR DAVID YAÑEZ COLIN

ASESOR:

DR. PEDRO ENRIQUE GARCIA RUIZ



CIUDAD DE MÉXICO

MAYO 2018



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

INTRODUCCIÓN

Muchas personas expertas en distintas disciplinas comparten la opinión de que la conciencia es el último santuario inmaculado que posee la humanidad, un espectáculo de magia y misterio maravilloso que además es capaz de albergar la naturaleza de sus propias almas y ser la habitación de todo lo que da sentido a la vida; algo en lo que es necesario invertir muchos recursos y procurar comportamientos dignos de merecer elogios y recompensas más allá de la vida terrenal. Esa idea no es errónea del todo, la conciencia es una de las verdades fundamentales de la existencia del ser humano, todos y cada uno de nosotros somos conscientes, somos capaces de acceder sin intermediarios al flujo de nuestra vida interior y al hacerlo, experimentar el amor, la satisfacción, el placer y el dolor, la felicidad, los recuerdos y anhelos, la imaginación, el futuro, la comprensión de ideas brillantes, la responsabilidad moral, la reflexión, la fe, la libertad, etc. La conciencia es la razón de vivir, si no fuéramos conscientes, nada en nuestras vidas tendría sentido o valor. Pero al mismo tiempo nada de lo que sabemos acerca del universo nos da suficientes información sobre qué es la conciencia ni cuál es su lugar en el orden natural, incluso hasta hacernos pensar que su naturaleza desafía toda explicación racional. No parece que la mente consciente pueda ser simplemente un objeto físico, alguna de sus partes o alguna de sus dinámicas. Aun cuando lo más obvio para un individuo consciente sea su propia mente, sigue teniendo cabida la pregunta de cómo es posible que los cuerpos físicos que habitan el mundo físico sean capaces de producir tales experiencias.

Aquellas personas que están tan comprometidas en defender el aspecto maravilloso de la conciencia también estarían de acuerdo en que algo como una Ciencia de la Conciencia es enteramente imposible. Esto se debe a que la ciencia es por naturaleza y convicción objetiva, mientras que la conciencia es por naturaleza subjetiva; y es por ser subjetiva que es tan valiosa; según esta perspectiva, resultaría en ocasiones conveniente que los esfuerzos de las explicaciones objetivas fracasaran, sería un atentado intelectual desmitificar el fenómeno de la conciencia. Pero es sencillamente una falta de correspondencia sostener que la comprensión de la conciencia excede el entendimiento humano hasta afirmar que no se trata de un problema sino un misterio, la conciencia no es misteriosa, tenemos una comprensión más o menos amplia de fenómenos físicos y

biológicos, como la reproducción, la auto-reparación celular, el origen del tiempo, la evolución de las especies, la deriva continental, la transmisión hereditaria de caracteres físicos, de las alteraciones del comportamiento bajo sustancias químicas, etc. Experiencias que en algún momento de la historia de la humanidad resultaban misteriosas y que aún hoy en día siguen asombrando y arrojando detalles a las comunidades intelectuales más vanguardistas, pero sin ser maravillosas, misteriosas o inexplicables en absoluto. Así como se han explicado estructuralmente estos misterios de la naturaleza, es esperable y deseable que la conciencia, en tanto verdad fundamental del ser humano, sea explicada con ese detalle y determinar de qué modo está contenida en el orden natural, quizá mediante una ciencia que enmarque sistemáticamente todos los aspectos involucrados.

La conciencia no es simplemente un interruptor de encendido/apagado, tiene una estructura compleja, en ella se manifiestan y participan fenómenos subjetivos y dinámicas objetivas. Un organismo que ha evolucionado hasta la conciencia posee estructuras biológicas, físicas y químicas objetivas correlacionadas con comportamientos y otras experiencias subjetivas como la percepción, las sensaciones corporales, las experiencias emocionales y una corriente de pensamientos de los que se da cuenta y también algunos de los que no se da cuenta. Ambos componentes aun necesitan explicación, tanto las estructuras y dinámicas físicas de los sistemas conscientes, como las experiencias subjetivas asociadas con ellas, pero en esas condiciones es posible generar un marco de explicación de las correlaciones que existen entre ellas. Describir estas dos dimensiones y su posible correlación es lo que nos hemos propuesto realizar aquí.

El capítulo I está dedicado enteramente a describir la naturaleza de la experiencia subjetiva. En la sección uno, describimos algunos tipos de manifestaciones de la experiencia subjetiva y sus cuatro características principales; que son la razón de que se considere irremediamente inexplicable por medios objetivos e identificamos estas características con el concepto de Qualia. Apuntalamos la idea de la naturaleza subjetiva de la experiencia apoyándonos en algunos argumentos epistemológicos. En la segunda sección consideramos la perspectiva de la primera persona como la manera más fiel de identificar la subjetividad de la conciencia. En la tercera sección hacemos una introducción a los datos en primera persona de los fenómenos mentales como contenedores de la experiencia subjetiva y describimos su cualidad representacional mediante el concepto de intencionalidad. En la

sección cuatro problematizamos la necesidad de explicar los datos subjetivos representacionales en términos objetivos, así como una perspectiva que muestra que podría ser imposible tal reducción. Introducimos un esquema general para la reducción llamado superveniencia y lo contrastamos con la dependencia causal entre sucesos mentales y hechos físicos, para mostrar que la experiencia subjetiva no puede ser reducida a explicaciones objetivas. Consideramos de nuevo la perspectiva de la primera persona para mostrar que la manera en que los sistemas conscientes acceden al mundo es en primera instancia de un modo subjetivo y concluimos que la experiencia subjetiva posee un lugar en el orden natural que debe ser tomado con seriedad. Si es que deseamos construir una Ciencia de la Conciencia completa, no podemos dejar fuera los aspectos subjetivos, aun cuando no comprendamos del todo su naturaleza. Finalmente en la sección cinco describimos las características generales de las posturas filosóficas dualistas sobre las cuales se construyeron los argumentos de todo el capítulo I, las cuáles apoyan la existencia de la subjetividad afirmando que la inteligencia consciente reside en algo no-físico.

En el capítulo II presentamos cuatro posiciones filosóficas que buscan refutar las ideas tradicionales dualistas y concentrarse en la objetividad de las estructuras y dinámicas objetivas. Estas cuatro posiciones están enmarcadas en el materialismo y muestran aportaciones para la explicación de la naturaleza de la conciencia tan relevantes que es de ingente importancia considerarlas. En la sección uno desarrollamos los principales argumentos del conductismo filosófico cuya tesis central es que la variedad de experiencias subjetivas que popularmente son reconocidas como estados mentales, en realidad no existen sino que no son más que una manera de designar la conducta de las personas. Analizamos la postura lingüística del positivismo lógico para la conducta y la posible elaboración de una teoría psicológica comprobable empíricamente y la reducción de las proposiciones de la psicología a proposiciones sobre conductas visibles. Y finalmente formulamos la argumentación acerca de las disposiciones de conducta que son consideradas por el conductismo como el material que da cuenta de la explicación de los estados mentales. Concluimos el capítulo señalando que el marco teórico del conductismo, presenta el primer avance materialista de la explicación sobre la conciencia, sin embargo considerar a la experiencia subjetiva como un elemento irrelevante contribuye con las principales críticas que recibió.

En la sección dos revisamos la más simple de las afirmaciones materialistas, la teoría de la identidad, cuyo principal argumento es: los estados mentales *son* estados físicos en el cerebro. Desglosamos la argumentación mediante el concepto de reducción interteórica y transformamos la discusión mente-cuerpo en la discusión mente-cerebro. Si es verdad que la conciencia es algo que ocurre en el cerebro entonces se debe desentrañar el papel causal que tiene su estructura o dinámica en función de dar como resultado el tipo de fenómenos de los que habla la experiencia subjetiva, para averiguarlo, revisamos algunos argumentos sobre la causalidad a favor de la teoría de la identidad. Finalmente concluimos que la teoría de la identidad es una hipótesis científica razonable y optimista con respecto a los avances de otras áreas de investigación sobre las funciones cerebrales, y apuesta a que en el futuro tales investigaciones arrojarán luz sobre las discusiones filosóficas de la conciencia.

En la sección tres revisamos el funcionalismo, la tesis que afirma que si una propiedad funcional pertenece a un objeto es exclusivamente debido a la aptitud que tiene el objeto de cumplir con cierto papel causal en un determinado contexto. Claramente es una tesis que surgió como complemento de la teoría de la identidad pero que poco a poco adquirió la relevancia única que la ha hecho convertirse en una de las teorías más importantes sobre la conciencia y sobre otras áreas de la ciencia. En esta sección revisamos el funcionalismo de maquina de Turing como la versión más efectiva de funcionalismo y el principio de realizabilidad múltiple, según el cual no interesa el material del que esté constituido un sistema consciente si efectúa la economía funcional causal de un estado mental. Y hacemos una breve revisión de la inteligencia artificial que es lógicamente compatible con el funcionalismo.

Finalmente en la sección cuatro presentamos la teoría materialista más radical, el materialismo eliminativo y lo acompañamos con la intuición sobre el naturalismo mecanicista. Esta primer postura es la tesis de que nuestra concepción de sentido común acerca de los fenómenos psicológicos que contienen experiencia subjetiva constituyen una teoría radicalmente falsa, una teoría tan defectuosa que tanto sus principios como su ontología serán desplazados por una neurociencia completa, en lugar de ser identificados con las conductas, las estructuras o dinámicas cerebrales o de sus papeles causales en sus determinados contextos. Para argumentar a favor de esta teoría hacemos una breve

descripción del concepto de psicología popular, intencionalidad y qualia desde una perspectiva crítica. Argumentamos de acuerdo con ciertos generadores de intuición acerca de los errores populares que cometemos cuando consideramos nuestro propios qualia. Y finalmente transportamos toda esa intuición hacia el plano no filosófico en términos de la tesis del naturalismo mecanicista para intentar encontrar el lugar natural de la mente una vez descritas sus características materialistas, sus funciones y su papel representacional y depurado las limitaciones que encuentra si la consideramos desde una posición dualista.

Concluimos el trabajo haciendo un repaso de la tradición experimental que considera estas dos dimensiones de la conciencia y que en la práctica ha descubierto que los experimentos suelen diseñarse en términos mecánicos y naturalistas pero siguen conteniendo aspectos subjetivos que arrojan riqueza a las investigaciones. ¿Cómo podemos construir una ciencia de la conciencia? En primer lugar debe poder explicarse en los mismos límites científicos en los que se explican otros fenómenos y otras energías. En última instancia una ciencia de la conciencia tendrá la tarea de integrar sistemáticamente las dos clases de datos que se describen en este trabajo en un mismo marco científico; la experiencia subjetiva y las estructuras o dinámicas naturales.

CAPITULO I

LA NATURALEZA DE LOS DATOS EN PRIMERA PERSONA

I.1. La experiencia subjetiva

Presumiblemente un estado mental auténtico está compuesto de dos partes, el proceso físico que lo origina (dinámica y estructura físicas) y la experiencia subjetiva particular asociada a este funcionamiento. Concentrémonos ahora en la segunda. Cuando hablamos de experiencia podemos referirnos a varias acepciones del concepto. Podemos decir que alguien tiene mucha experiencia en el diseño editorial, o en el diagnóstico médico; en estos casos hablamos de que alguien tiene mucha práctica en ciertas capacidades o habilidades y que es capaz de llevarlas a cabo con gran eficiencia, por ejemplo a la luz de un imprevisto, porque lleva mucho tiempo haciéndolo. También podemos referirnos a la experiencia como un suceso que ha marcado nuestras vidas, decimos de ese suceso que fue una experiencia inolvidable. Sin embargo, ninguna de esas dos acepciones nos interesa aquí. Lo que nos interesa en general son los sucesos o estados mentales en su forma más básica; la primer capa de receptividad desde el mundo *exterior* hacia el *interior* de nosotros mismos, es decir, aquellos sucesos que intervienen en la percepción, en la sensación y en la reflexión. En particular nos interesa la subjetividad asociada a esos sucesos. En este capítulo nos centraremos en la naturaleza de esa parte subjetiva de los estados mentales.

Pensemos lo siguiente: todo lo real posee propiedades, la experiencia consciente es real, así que ésta posee también propiedades. Los estados de conciencia de cada persona poseen ciertas propiedades en virtud de las cuales esos estados (mentales) tienen las características que tienen. Esto significa que siempre que un sujeto experimenta algo de una manera, y no de otra, de modo que tal experiencia adquiere un carácter subjetivo, esto es verdad en virtud de alguna propiedad de algo que ocurre en esa persona, en ese momento particular. Con más claridad lo que queremos decir es que lo esencial de la experiencia subjetiva es que cuando alguien experimenta algo, ese algo se ve, se escucha, se siente, sabe o huele de cierta manera; quizá diferente de como esa experiencia es realmente, presentada de modo adecuado a las condiciones del entorno y a las capacidades que el sujeto posea para ser estimulado por esa experiencia, o incluso, diferente del modo en que

se disponga luego a comportarse o a creer en ello. Es decir, existen sabores, experiencias táctiles, aspectos visuales, aromas y sonidos subjetivos que se nos presentan de una manera determinada, que bien podrían ser *diferentes* cuando los experimentamos los humanos, del modo en que lo experimentan otros organismos, y más aún, diferentes totalmente de la manera en la que son en realidad,¹ si es que tienen una manera de ser independiente del modo en que se presentan ante las mentes capaces de percibirlos.²

Según la visión tradicional de la experiencia subjetiva, ningún sujeto es capaz, ni siquiera de sugerirle a otro, por creativo y elocuente que pueda ser y por imaginativo y cooperativo que resulte su interlocutor, la manera exacta en la que experimenta algo de manera que tal experiencia es irremediabilmente inefable.³ Tal característica se debe a que las propiedades de las que hablamos son intrínsecas; así el sujeto puede estar completamente seguro de las cualidades que su propia experiencia posee, ya que se le presentan de manera directa en la conciencia. Hablamos entonces de que la experiencia subjetiva posee cuatro características básicas: es intrínseca, inefable, privada y directamente aprehensible por el sujeto que la experimente.

La experiencia subjetiva ha sido identificada con diferentes acepciones para señalar tales características: propiedades introspectivas de la sensación, conciencia fenoménica, carácter cualitativo, contenido intrínseco o simplemente *qualia*.⁴ Estos términos filosóficos tienen la finalidad de englobar la referencia subjetiva que poseemos los humanos con respecto a las propiedades de las cosas del mundo externo que nos afectan sensorialmente o

¹ Cualesquiera que sean las propiedades informacionales, disposicionales o funcionales que posea el sujeto, ninguna de ellas será capaz de advertir la parte subjetiva de las experiencias, según la visión tradicional. Estos puntos serán discutidos con mayor profundidad en el capítulo II.

² Nos referimos aquí a la dicotomía realismo/idealismo. Mientras lo primero afirma que las propiedades de los objetos son independientes de la mente, es decir, objetivos, lo segundo afirma que si no existe una determinada mente que perciba tales propiedades en los objetos, entonces esas propiedades no existen, de manera que las propiedades de los objetos son dependientes de las mentes, es decir, subjetivas.

³ No solo se descarta la comunicabilidad verbal, también las pruebas experimentales, fisiológicas o conductuales que intentan trasladar la experiencia subjetiva al campo objetivo.

⁴ En la *Stanford Encyclopedia of Philosophy* (puede consultarse en línea en: <http://plato.stanford.edu/entries/qualia>) Michael Tye presenta el concepto de la siguiente manera: Los sentimientos y las experiencias pueden ser muy variados: paso la mano por un papel de lija, huelo un zorrillo, siento un dolor agudo en la mano, creo ver algo de un intenso color púrpura, me enojo mucho. En cada uno de estos casos, soy el sujeto de un estado mental con características distintivas muy claras. Lo que se siente al protagonizar cada estado es específico, tiene una fenomenología determinada. Los filósofos utilizan el término “qualia” (en singular “quale”) para referirse a los aspectos fenoménicos de nuestra vida mental, accesible por medio de la introspección.

sobre nuestras disposiciones de conducta o creencia. En otras palabras, son términos muy rebuscados para lo que presumiblemente no podría resultar más directamente accesible a cualquier persona: la experiencia subjetiva es la manera en que nos parecen las cosas.

¿Qué tipo de experiencias conscientes son subjetivas, es decir, qué experiencias poseen tales cualidades intrínsecas? A continuación se presentan seis ejemplos de experiencias subjetivas asociadas a estados mentales:

1. Experiencias visuales

Consideremos una experiencia visual como el estado mental al que alguien está sometido cuando aprecia el color azul turquesa en el catálogo de una tienda de pinturas. Hay algo que es para él el hecho subjetivo de experimentar el azul turquesa; lo que es muy diferente del hecho subjetivo de experimentar, por ejemplo, el amarillo “Kandinsky” en el mismo catálogo. Ambos colores se presentan ante su sentido de la vista de una manera particular, cualitativamente distinta entre sí, de manera que es capaz de reconocer la diferencia entre uno y otro. Cualquier persona que se sometiera al mismo ejercicio, podría sin ningún problema detectar la diferencia entre los dos, gracias a que los colores se le presentarán también como un hecho subjetivo cualitativamente distinto.

2. Otras experiencias perceptuales

Cualquiera puede elegir una sonata de Beethoven, escucharla con atención es una experiencia *sensacional*, puede agregarle la propiedad de ser bella o aburrida, darse cuenta de que sus compases son armónicos y luminosos, que su ritmo es contagioso. Puede salir al jardín y escuchar el canto de las aves, el crujir de las hojas de los árboles y tirarse en el pasto y sentir su frescura. En todas esas experiencias esa persona es el protagonista de una serie de fenómenos subjetivos acerca de su manera de darse cuenta de las cosas que ocurren en el mundo. También están presentes nuestras experiencias de lo dulce y lo salado, lo ácido y lo amargo, así como nuestra capacidad de “saborear con la nariz”; una experiencia indescriptible pero detectable.

3. Experiencias corporales

Una persona ha recibido un fuerte golpe en la espalda, lo que le ha provocado una lesión en el nervio radial de su brazo izquierdo. Esa es la causa de que las siguientes semanas experimente un terrible dolor en la muñeca. Acude al médico y explica que el dolor es muy agudo, parecido a un ardor, sordo y difuso, acompañado de una punzada que recorre todo su brazo hacia arriba hasta la base del cuello, que le hace experimentar sensaciones aberrantes, desconocidas y difíciles de describir. ‘Agudo’, ‘ardor’ ‘sordo’, ‘difuso’ y ‘aberrante’ son palabras con las que intenta señalar su experiencia subjetiva de dolor; gracias a las cuales el médico es capaz de anticipar el tipo de lesión que ha sufrido, pero no gracias a las cuales puede sentir la sensación dolorosa del paciente.

4. Imágenes mentales

Alguien se despierta en la mañana con el recuerdo un sueño. En él observó una orquesta vestida de rojo que recorría grandes jardineras al atardecer, tocando música de carnaval. En este caso ni el sueño ni el recuerdo de él están constituidos por elementos objetivos, es decir, ninguno de los integrantes de la orquesta poseía, de hecho, un sombrero cuyo color rojo fuera accesible a cualquier individuo, por el contrario si esa persona no lo hubiera soñado rojo, no sería rojo (cabría en este punto preguntarnos ¿qué objeto es rojo?); además de que solo esa persona lo ha soñado. Existe la posibilidad de que los sombreros fueran amarillos, de manera que solo el individuo que lo soñó es el responsable de las propiedades de los objetos en el sueño. Lo mismo para las propiedades de los objetos que describe en su recuerdo; quizá aunque no hubiera puesto suficiente atención al color de las plantas en las jardineras, al relatar el sueño, las describa verdes, y así sucesivamente con cada elemento del sueño y del recuerdo del sueño. En esta situación esa persona es de nuevo el protagonista de experiencias subjetivas, con una fenomenología determinada dependiente de su manera de percibir las, de la manera en que aparecen en su *interior*. Si le relatará el sueño a alguien más, su interlocutor no tendría otra opción más que confiar en que en el sueño los sombreros eran rojos y la música de carnaval.

5. Experiencias emocionales

Una pareja de novios en la biblioteca ha hecho llegar al límite de su paciencia a la chica que tienen al lado. Se besan y hacen ruido, mientras ella intenta concentrarse. Por su manera de mirarlos, seguro está a punto de explotar furiosa para callarlos. Por fin ocurre para beneplácito de todos los demás presentes. Se oyen por toda la sala frases de apoyo, risas y quejas. En este caso ha habido distintas respuestas, derivadas de la experiencia particular que cada individuo experimentó ante el evento; enojo, alegría, contento, etc.

6. Pensamientos ocurrentes

En una de las paredes del salón se exhibe un enorme letrero con la leyenda “Prohibido fumar”. Existe un acuerdo intersubjetivo entre los presentes tal que a nadie se le ocurriría jamás fumar adentro. No obstante al fondo alguien ha decidido romper la regla. -¿Qué pasa por su cabeza?- Piensan los demás desconcertados. Algo diferente debe estar ocurriendo, de manera que a él no le parece importante ni el letrero ni el acuerdo intersubjetivo del que participan todos. Él opina que no hay ningún problema en fumar dentro del salón, aunque esté prohibido hacerlo.

La clase de fenómenos de los ejemplos anteriores ocurren en los sujetos capaces de darse cuenta de ellos, así que es de suponerse que cualquier ser humano en condiciones ambientales y de salud física normales, es susceptible de tenerlos. El problema al que se enfrentan es que parece que el estudio de los datos que podrían ser extraídos de ellos, exige necesariamente una perspectiva objetiva si lo que se pretende es alcanzar un modelo explicativo de la realidad completo (e. g. el estudio comprobable de los fenómenos físicos que pudieran originarlos o de aquellas conductas que revelen su presencia). Hablamos, entonces de que sería necesario estudiar *datos* de fenómenos con características esencialmente subjetivas, mediante una perspectiva que encaje con un modelo de realidad objetivo; y esto supone al mismo tiempo que la experiencia subjetiva de los ejemplos anteriores, no es la clase de cosas que encajan en ella. Aunque por cierto, sea muy fácil reproducir estos ejemplos tomando casos de la vida cotidiana.

La idea tradicional de la experiencia subjetiva puede rastrearse desde Sexto Empírico (circa 160-c. 210). Pero no fue sino hasta el desarrollo del pensamiento de autores como René Descartes y posteriormente John Locke y Gottfried Leibniz, que adquiere las dimensiones que darían pie a las discusiones contemporáneas. Analicemos a continuación las ideas centrales de estos autores.

René Descartes en las *Meditaciones* (Descartes R. , 1641/2003), buscaba encontrar cualquier conocimiento que fuera absolutamente cierto, del cual no existiera ni la menor duda. Para empezar dudaba de todo cuerpo material, incluso de su propio cuerpo. Se preguntaba si sería tan dependiente de su cuerpo o de sus sentidos, de manera tal que sin ellos no pudiera ser nada. El hecho es que si había llegado a persuadirse o simplemente llegado a pensar algo, es sin duda, porque él era algo. Aún en el caso de que hubiera un engañador astuto que intentara engañarle sobre todas las cosas que conocía, por mucho que lo intentara, no podría lograr que él pensara que no es algo. De manera que la proposición “yo soy, yo existo”, o cualquier otra, es completamente verdadera mientras Descartes *piense* que existe, (*Cogito ergo sum*: Pienso, por lo tanto existo).

¿Exactamente qué es Descartes? había considerado hasta entonces que tenía un rostro, manos, brazos, piernas, y en general todo aquello que constituye toda la maquinaria física compuesta de huesos y carne (nervios y funcionamiento determinado) que llamaba cuerpo. Además, se alimentaba, andaba, sentía y pensaba; actividades que le atribuía al *alma*. Si le hubieran pedido que definiera al cuerpo, él habría dicho que es todo aquello que está delimitado por una forma, que está colocado en cierto lugar y que llena un espacio de modo que excluya a cualquier otro cuerpo; todo aquello que pueda ser sentido por el tacto o por la vista o por cualquiera de los cinco sentidos básicos, que pueda moverse de varias maneras, no por sí mismo sino por algún objeto externo a él que le imprima alguna fuerza. Ya que no creía posible que el cuerpo se moviera, sintiera o pensara por sí mismo. No obstante de todas esas cosas era posible dudar hasta el momento (recordemos la duda metódica), ya que había postulado la hipótesis de la existencia de cierto genio maligno que podía engañarlo sistemáticamente, así que era posible que nada de eso tuviera una existencia cierta. Por tal motivo era necesario que examinara ahora los atributos de su *alma* para ver si en ellos podía encontrar la certeza que buscaba sobre qué era él.

Alimentarse y andar eran actividades del alma pero no cosas que pudieran ser ciertas, ya que si es posible que no tuviera cuerpo, entonces no habría nada que necesitara ser alimentado, ni nada que anduviera. Tampoco sentir, ya que aquello que siente es el cuerpo; cabe mencionar que las sensaciones caen presa de otra de las hipótesis escépticas, la del sueño: en sueños también hemos experimentado diversos escenarios mentales, de manera que ¿cómo identificamos que al sentir o percibir, no estamos soñando?

Pensar era la última actividad que le atribuía al alma y encontró que el pensamiento es aquello que no puede separarse de su propia existencia. “Yo soy, yo existo” es cierto mientras Descartes piense; de hecho, mientras piense esa o cualquier otra proposición. De manera que Descartes era para entonces sólo una cosa que piensa. En términos actuales, una mente. Una mente o una cosa pensante es una sustancia no física que reflexiona, que percibe y que siente, que experimenta ciertas cosas como por los sentidos, ve luz, oye ruidos y siente el calor, cosas que podrían ser falsas pues existe la posibilidad de estar soñando; lo que no puede ser falso es que a él le parece sentir calor u oír ruido, lo que no podría suceder si no estuviera sintiéndolo, es decir, si no estuviera teniendo una determinada actividad mental. Le ha quedado claro después de examinar las cualidades primarias y secundarias de un pedazo de cera que las cualidades de los cuerpos no son propiamente conocidos por los sentidos o por la facultad de imaginar, sino por el entendimiento sólo; no son conocidos porque los vemos y tocamos, sino porque los entendemos o comprendemos por el pensamiento.

Los cuerpos que tocamos y vemos son los objetos que conocemos con mayor facilidad, los que conocemos, según Descartes, más distintamente. En un pedazo de cera, conocemos su dulzura, su olor, su color, su figura, su dureza, su ‘frialidad’ y cosas por el estilo, tales cosas constituyen lo requerido para que un cuerpo pueda ser conocido lo mejor posible. Pero al ser calentada, todas esas propiedades cambian, incluso algunas desaparecen, como su olor, su dureza, su figura y su sonido al ser golpeada. Nadie negaría que aun cuando esas propiedades han cambiado o desaparecido, permanece la misma cera, ¿pero qué se supone que era aquello que conocíamos con más facilidad? No puede ser ninguna de las cosas que se percibían con los sentidos, pues esas cualidades han cambiado o desaparecido. Lo que queda si le restamos esas cualidades, es un objeto extenso, que puede modificar su forma en variedades inimaginables. Por lo tanto, la cera (o la parte de la cera) que sólo puede ser

percibida con la mente es pues, la misma que se ve, que se siente, que emite un sonido al ser golpeada, etc. Aquellas cosas que constituyen sus cualidades secundarias, aquellas que no son nada en los objetos mismos, sino funciones que producen en nosotros diversas experiencias por medio de sus cualidades primarias; que producen, de hecho, experiencias subjetivas.

En el *Ensayo sobre el entendimiento humano* (Locke, 1690/1999) John Locke postula la existencia de dos tipos de cualidades en los objetos, a partir de las cuales se generan ideas en la mente. Una idea según Locke puede ser todo aquello inmediato a la sensación o a la percepción en nuestro entendimiento. Una cualidad de un objeto, en general, es el poder de producir en la mente una idea. Los dos tipos de cualidades de los objetos son: las cualidades primarias y las cualidades secundarias. Las cualidades primarias son aquellas enteramente inseparables del cuerpo, cualquiera que sea el estado en el que se encuentre; en otras palabras, todas aquellas propiedades objetivas, aquellas que pueden ser captadas por los sentidos en cada partícula del objeto, tales como la extensión, la forma, el movimiento, el reposo y el número. Las cualidades secundarias, entonces, son aquellas que tienen, según Locke, el poder de producir en nosotros determinadas sensaciones, a saber, el color, el sonido, el sabor, el olor y la experiencia táctil. Incluso en el mismo orden de ideas, podrían producir en nosotros deseos, reflexiones o creencias. La experiencia particular que provoca en nosotros depende de las cualidades primarias que podamos percibir en el objeto, pero en ningún sentido podríamos decir que esas ideas se le parecen a alguna cualidad perteneciente al objeto en sí mismo, es decir, no existe nada en los objetos que se asemeje a nuestro deseo. Las ideas derivadas de ambas cualidades tienen la siguiente forma:

§ 15. Las ideas de las cualidades primarias de los cuerpos son semejanzas de dichas cualidades, y que sus modelos realmente existen en los cuerpos mismos; pero que las ideas producidas en nosotros por las cualidades secundarias en nada se les asemejan. Nada hay que exista en los cuerpos mismos que se asemeje a esas ideas nuestras. En los cuerpos a los que denominamos de conformidad con esas ideas, sólo son un poder para producir en nosotros esas sensaciones; y lo que en idea es dulce, azul o caliente, no es, en los cuerpos que así llamamos, sino cierto volumen, forma y movimiento de las partes insensibles de los cuerpos mismos; pero que en nada se asemejan las ideas que en nosotros producen las cualidades secundarias. (Locke, 1690/1999, pág. 51)

En otro contexto histórico, Leibniz escribía en la *Monadología*:

Hay que reconocer, por otra parte, que la percepción y lo que de ella depende resultan inexplicables por razones mecánicas, esto es, por medio de las figuras y de los movimientos. Porque, imaginemos que haya una máquina cuya estructura la haga pensar, sentir y tener percepción; se la podrá concebir agrandada, conservando las mismas proporciones, de tal manera que podamos entrar en ella como en un molino. Esto supuesto, una vez dentro, no hallaremos sino piezas que se impulsan unas a otras, pero nunca nada con que explicar una percepción. Así, pues, esto hay que buscarlo en la sustancia simple, no en el compuesto o en la máquina. Más aún, no cabe hallar en la sustancia simple otra cosa excepto esto, es decir, excepto las percepciones y sus cambios. Y solamente en esto también pueden consistir todas las acciones internas de las sustancias simples. (Leibniz, 1721/1981, pág. 87)

La preocupación de esta tradición es que la existencia de las cualidades secundarias sólo es reconocible en el ámbito de lo mental. Los cuerpos mismos no contienen ninguna de esas cualidades, ellas existen únicamente como propiedades de algo en las mentes capaces de detectarlas. Si quisiéramos estudiar nuestras propias experiencias subjetivas del tipo antes descrito de acuerdo con las ideas de esta tradición, únicamente debemos ser capaces de “echar un ojo dentro de nosotros mismos”, dentro de nuestras propias mentes. Es decir, no sólo podemos mirar el color amarillo “Kandinsky” en el catálogo de pinturas con nuestros ojos o escudriñar el olor de una flor con nuestra nariz en el jardín; también podemos, sin necesidad de los órganos de los sentidos observar algo extra cuando ponemos suficiente atención: la corriente de experiencias subjetivas en nuestro interior. Esto quiere decir que podemos llevar a cabo una introspección hacia nuestra propia conciencia para averiguar qué experiencias subjetivas ocurren en ella.

Cuando llevamos a cabo una introspección, es decir, cuando fijamos la atención en el contenido de nuestra conciencia, nos damos cuenta del flujo de pensamientos, deseos, creencias, percepciones, sensaciones y emociones. Si es correcto que los estados mentales están compuestos de experiencias subjetivas y procesos físicos, la introspección no es algo que nos ayude a descubrir la naturaleza de éste último (*e.g.* la incorrecta actividad de pulsaciones electroquímicas en el sistema nervioso periférico causante de la neuropatía del ejemplo 2) En cambio nos permite percatarnos de que el dolor es punzante, agudo, sordo y difuso, características que tales conceptos anticipan como una experiencia totalmente subjetiva. Por lo tanto la introspección revela la naturaleza subjetiva de los estados mentales, no el proceso físico.

Se podría acusar a esta forma de presentar las cosas como una mera estrategia de persuasión, sin articular argumentos concluyentes sobre ello esgrimiendo una retórica del tipo: “¡nadie puede negar que existen tales cosas! Todos hemos recordado un sueño o hemos tenido hambre y nadie puede negar que asociado a cualquiera de esos eventos ocurre algo que no se puede compartir con nadie más, sólo nosotros sentimos nuestra hambre, sabemos a qué nos sabe la miel y tenemos acceso a nuestros propios sueños y recuerdos”. Pero no se trata aquí de apelar a una falacia *ad populum* y conformarnos con que sea cierto que tanto a mí como a ti, la miel nos sabe de una forma, a la que yo no tengo acceso si es *tuya*, ni tú tienes acceso si es *mía*. Existen argumentos firmes que pretenden apoyar la realidad de la experiencia subjetiva. Consideremos los siguientes tres.

El primero de ellos se conoce como el “argumento del conocimiento”. Este argumento fue propuesto por Frank Jackson en su artículo “Qualia epifenómicos” (Jackson, 2012) y extendido en su artículo “Lo que María no sabía” (Jackson, 2012). En términos generales nuestra interpretación de la idea central del argumento sugiere que:

1. La experiencia subjetiva existe y que ninguna explicación objetiva, por más sofisticada que sea, puede explicarla.
2. En la experiencia consciente participan, además de los hechos que pueden ser estudiados por la física, hechos no físicos.
3. Y apoya en última instancia el hecho de que un estado mental auténtico está compuesto por un evento físico y una parte no física, es decir una experiencia subjetiva asociada a un evento físico; que de hecho no puede ser anticipada, ni predicha, ni explicada por éste. Pero que aún en esas condiciones está presente y que acceder a ella representa un aumento en el conocimiento, diferente del conocimiento que proporciona el estudio de los hechos físicos.

El argumento tiene dos versiones:

La primera versión es la siguiente: Alfredo tiene la mejor visión del color que cualquiera de los que se tiene registro en la humanidad; ha hecho todas las discriminaciones de color nunca antes hechas y además hace otra que el resto ni siquiera podemos anticipar.

Si se le muestra un montón de tomates maduros los clasificará en dos grupos con absoluta consistencia, es decir que si se le pide que los clasifique de nuevo, después de vendarle los ojos y mezclar los tomates, lo hará exactamente de la misma forma que la primera vez.

Alfredo nos explica que no todos los tomates se ven de la misma forma, él ve dos colores en donde los demás vemos sólo uno, y en consecuencia ha desarrollado para su propio uso los términos rojo 1 y rojo 2 con el fin de identificar la diferencia. Ha intentado enseñar la diferencia a sus amigos, sin éxito, y ha llegado a la conclusión de que el resto del mundo es daltónico con respecto a los colores rojo 1 y rojo 2 que él percibe; así que usa el término común “rojo” para adecuarse al uso convencional. Rojo 1 y rojo 2 son colores cualitativamente diferentes para él, no es simplemente una diferencia en la variación de tonalidades del rojo común, sino que son tan diferentes como lo serían para nosotros el amarillo del morado, en sus distintas variaciones de tonalidades.

Supongamos que hacemos varios estudios físicos de Alfredo. Una investigación de la base fisiológica de su capacidad revela que su sistema óptico es capaz de separar dos grupos de longitudes de onda en el espectro del rojo, con tanta nitidez como los demás separamos las longitudes de onda del amarillo y el morado. Otros estudios descubren que los conos de Alfredo responden diferencialmente a ciertas ondas luminosas en la sección del espectro del rojo que en el resto de las personas no hace ninguna diferencia. Pero hay un problema, estos estudios no nos dicen lo que en realidad queremos saber, ¿qué clase de experiencia tiene Alfredo cuando ve el rojo 1 y el rojo 2? ¿Cómo son esos dos colores y en qué se diferencian de nuestro limitado y normal rojo? Sin duda hay algo acerca de ellos que no sabemos, incluso en el caso de que supiéramos todo lo relacionado con la teoría del espectro visual humano (siempre y cuando ésta fuera totalmente irrefutable) y la especial fisiología de Alfredo.

Imaginemos que como resultado de la investigación, logramos modificar quirúrgicamente la fisiología de cualquier persona para que sea igual a la de Alfredo (después de todo sabíamos que había algo diferente en la constitución física de Alfredo, era de esperarse que modificando lo físico pudiéramos modificar también el acceso a la experiencia subjetiva); se presentaría entonces la posibilidad de saber al fin como es ver otro color, y al mismo tiempo descubrir la diferencia entre la visión normal y la visión extendida que poseía Alfredo. Digámoslo de la siguiente manera, antes de la operación

teníamos toda la información física posible para comprender la experiencia subjetiva extendida del color rojo de Alfredo, pero eso nos dejaba con una incógnita (una enorme por cierto); después de la operación sabemos más acerca de Alfredo y en especial más acerca de la experiencia de color. Presentadas así las cosas podemos afirmar que la experiencia subjetiva está presente, sea o no explicada por los hechos físicos y que el acceso auténtico a ella proporciona un tipo de conocimiento diferente del que proporciona el estudio aislado de los hechos físico, aunque éste sea completo y cabal.

La segunda versión del argumento es la siguiente: María es una súper científica que ha sido forzada a estudiar el mundo, confinada a una habitación en blanco y negro y recibiendo toda su instrucción a través de un monitor también a blanco y negro, vestida con ropa y guantes a blanco y negro y demás detalles imaginativos a blanco y negro. Se especializa en neurofisiología de la visión y adquiere toda la información física que hay que obtener acerca de lo que pasa cuando vemos colores y cuando usamos los términos “rojo”, “morado” y demás. Conoce con precisión las combinaciones de longitudes de onda que estimulan en la retina los diversos colores, la función del cerebro para discriminarlos y exactamente como todo esto combinado con la contracción de las cuerdas vocales y la expulsión de aire en los pulmones produce oraciones como “Los tomates son rojos”.

Imaginemos que María es liberada de su confinamiento, sale de su habitación y por fin observa el sin fin de colores en el mundo. ¿Qué ocurriría? ¿Aprenderá algo nuevo o no? Parece perfectamente obvio que aprenderá algo acerca del mundo y de la experiencia visual que antes no sabía. Lo que supone que su conocimiento anterior era incompleto; no obstante suponíamos que era suficiente para saber todo acerca de los colores. Si es falso que la experiencia subjetiva existe, entonces, para saber todo acerca de la experiencia de color, sería suficiente saber todo acerca de los hechos físicos relacionados con ella, y que por supuesto María no aumentaría en ningún sentido sus habilidades al percibir los colores. Suponer esto es suponer que la experiencia subjetiva no enseña nada, porque ni siquiera existe. Por esa razón podemos concluir, hasta este punto que la experiencia subjetiva existe y proporciona información importante sobre la conciencia, al menos en lo relacionado con la percepción.

Hay que destacar que el propósito del argumento de Jackson es argüir que el fisicismo es falso dada su ineficacia para explicar la totalidad de los hechos presentes en la

conciencia, asegurando que aún si supiéramos absolutamente todo acerca de los procesos físicos relacionados con algunos estados mentales tan básicos como la percepción del color, eso no sería suficiente para explicar todo lo relacionado con tales estados mentales. Lo que quedaría sin explicar (o incluso negado, según la teoría desde la que revisemos el argumento) es justamente la experiencia subjetiva asociada a esos eventos físicos. Nuestra lectura es menos ambiciosa ya que sugiere únicamente la presencia de la experiencia subjetiva. En otras palabras, que en los estados mentales podemos identificar una parte física y algo más, interviniendo en la noción de conciencia. Tratamos, por el momento, de anticipar que el estudio del conjunto de los hechos físicos y la experiencia subjetiva, es el camino adecuado para explicar la conciencia, ya que si alguno de los dos es ignorado, parecería que el problema queda solo parcialmente explicado. Nos conformaremos con eso, más adelante profundizaremos en la ineficacia de los procesos físicos aislados para explicar la totalidad de la conciencia.

También debemos resaltar que a pesar de que nuestra lectura es más débil está comprometida con la versión más fuerte del argumento. Ya que lo que deseamos extraer de la conclusión es que realmente existen hechos no físicos que están relacionados con la visión humana y no sólo con la versión débil que afirma que hay un tipo de conocimiento que concierne a hechos acerca de la visión humana a color. Aunque el argumento está abierto a estas dos interpretaciones es obvio que Jackson está pensando en la versión fuerte y aquí concordamos en él.

La pregunta que subyace incluso si alguien se da cuenta completamente de todas las funciones objetivas involucradas en la conciencia es ¿Por qué todo ese funcionamiento está asociado con la experiencia subjetiva? Y más aún ¿Por qué está asociado con la particular clase de experiencia subjetiva con la que de hecho está asociado? Parece que simplemente explicando las funciones objetivas no respondemos esas preguntas.

El segundo argumento es conocido como el argumento modal. En su forma más general dice:

1. Es imaginable que la mente de un individuo pueda existir sin su cuerpo.
2. Es concebible que la mente de un individuo exista sin su cuerpo.
3. Es posible que la mente exista sin el cuerpo.

Por lo tanto:

4. La mente es una entidad diferente del cuerpo.

Este argumento ha sido ampliado a través de la posibilidad lógica de los zombis⁵. Esta pintoresca posibilidad nos dice que ninguna cantidad de información física acerca de otra persona (u organismo) implica lógicamente que él o ella sea consciente, o que sienta, perciba o reflexione algo en absoluto. Es posible concebir, en consecuencia un mundo en el que habitan organismos exactamente iguales a nosotros pero carentes en su totalidad de cualquier tipo de vida mental. Esto incluye cualquier tipo de estado funcional o historia física.

Un zombi es exactamente eso en la manera filosófica de hablar de ellos. Por ejemplo para David Chalmers un zombi (su gemelo zombi):

Es idéntico a mí, molécula por molécula, e idéntico en todas las propiedades inferiores postuladas por la física, pero carece por completo de conciencia [...] Está inmerso en un entorno igual al mío. Es idéntico a mí *desde el punto de vista funcional*; procesa el mismo tipo de información, responde a los estímulos del mismo modo que yo, con las modificaciones pertinentes de las mismas configuraciones internas y con consecuencias conductuales exactamente iguales [...] Está despierto, es capaz de informar acerca de sus estados internos y de centrar su atención en distintos lugares, y así sucesivamente. La única diferencia es que todo ese funcionamiento no va acompañado de ninguna experiencia consciente real. No tiene sensaciones fenoménicas. Para el zombi no se siente nada ser un zombi. (Chalmers, *La mente consciente: en busca de una teoría fundamental*, 1999, pág. 95)

Así que si hay una clase de sujetos idénticos a los seres humanos, ellos también discriminarían el azul turquesa del amarillo “Kandinsky”, también despertarían y relatarían un sueño, también opinarían que la sonata de Beethoven es aburrida, saldrían al jardín con el objetivo de sentir la frescura del pasto e incluso mostrarían que lo están disfrutando; pero ninguna de esas actividades tendría una sensación (percepción o reflexión) determinante en su forma de ser experimentada o sobre la disposición de conducta o creencia, o en términos

⁵ La existencia de un zombi es simplemente una posibilidad lógica, es decir cualquier prueba deductiva validaría su existencia sin ninguna contradicción.

de representación, etc. absolutamente todo dentro de ellos sería oscuro. Incluso Alfredo y María podrían ser zombis; no lo sabemos. Lo que sí sabemos es que en nosotros, los humanos, a diferencia de un zombi sí hay algo cualitativo en todas nuestras actividades; nuestra vida mental no sólo está compuesta de un proceso físico, tenemos algo más que eso: experiencia subjetiva.

Como para nosotros existen experiencias subjetivas que acompañan nuestras vidas (después de todo nosotros no sólo somos lógicamente posibles, de hecho existimos), podríamos decir que existe en general una manera de ser como nosotros mismos, en este orden de ideas, anticipábamos, no habría nada que sea ser como un zombi y sí algo que es ser como nosotros mismos. Este es el tercer argumento, que se diferencia del argumento del conocimiento en que aquel únicamente nos decía *algo acerca* de cierta cualidad de la experiencia. Es decir, en el caso de que desarrolláramos la habilidad quirúrgica de implantar en cada persona el mecanismo que le permitía a Alfredo reconocer dos colores en donde los demás sólo reconocemos uno, eso no sería equivalente a *ser* Alfredo. Quizá a él ver dos colores lo estresaba al no poder compartirlo, quizá a alguien más con la misma capacidad, le genere una idea para hacer fortuna; eso depende de la manera en que se dispongan los sujetos a comportarse ante su propia vida mental. Lo que en general quiere decir que ningún conocimiento acerca de nadie, sea físico o de cualquier otra índole, será conocimiento de ese sujeto “desde el interior”, porque nadie puede ser otro sujeto diferente del que de hecho es. Y eso nos deja en la posición de afirmar que esa imposibilidad de saber *cómo es ser* otro sujeto se debe a que cada quien posee experiencias subjetivas cualitativamente distintas, más allá de que obviamente es otro cuerpo físico.

La experiencia subjetiva como la hemos descrito anteriormente sólo pueden ocurrir en un ser capaz de darse cuenta de su entorno y de sí mismo. Simplemente no tiene sentido hablar de experiencias subjetivas de las que nadie se dé cuenta. Pero ¿Qué es ser consciente? ¿Qué significa que alguien sea consciente de una experiencia subjetiva?

I.2. La perspectiva de la primera persona: una manera de identificar la subjetividad

Hemos dicho que lo esencial de una experiencia subjetiva es ser *sentida* o experimentada por un sujeto consciente. Thomas Nagel en su artículo *¿Cómo es ser un murciélago?* (Nagel, 2012) Nos presenta la intuición que goza de mayor popularidad sobre qué significa ser un sujeto consciente. Como ya lo anticipábamos en la sección anterior se trata del “cómo es ser”.

Cualquier organismo tiene experiencias, un murciélago es un organismo, así que éste también tiene experiencias. Según Nagel, esas experiencias, son conscientes en la medida en que hay algo que es cómo se siente *la* experiencia de ser murciélago, para el murciélago, de modo que haya algo, para el murciélago, que sea ser como un murciélago. En estos términos que un organismo sea consciente quiere decir que, hay algo que es ser como ese organismo particular: para un murciélago ser como un murciélago, para un perro ser como un perro, para un humano ser como un humano o quizá para una máquina ser como una máquina, etc.

¿Podríamos nosotros saber cómo es ser un murciélago o cualquier organismo diferente del que somos? Lo que constituye la manera de ser de un organismo, de acuerdo con Nagel, son las experiencias que tiene y éstas están determinadas, al menos en parte, por los aparatos perceptuales de los organismos.⁶ La mayoría de las experiencias perceptuales de los murciélagos provienen principalmente de su aparato sonar, por medio de ecolocalización; si quisiéramos saber cómo es ser un murciélago, necesitaríamos saber cómo es tener una experiencia ecolocalizadora, y para saberlo necesitaríamos ser un murciélago. El problema es claro: no somos murciélagos, carecemos de un sentido perceptual semejante al sonar, y por ello también de las experiencias ecolocalizadoras necesarias para saber cómo es ser un murciélago.

⁶ Decimos que sólo en parte, porque debemos considerar que las percepciones no constituyen la totalidad de las experiencias subjetivas; existen también dentro de la misma categoría, las sensaciones de ardor, comezón o cosquilleo, que no parecen ser percepciones; y también están presentes los deseos y creencias asociadas a un hecho. Es importante señalar esta precisión, porque como veremos más adelante, incluir este tipo de experiencias en la categoría de experiencia subjetiva requiere de argumentos más precisos.

Si no podemos saberlo, entonces al menos ¿Podemos imaginar cómo es tener experiencias ecolocalizadoras para llegar a saber cómo es ser un murciélago? Podría imaginarme que tengo los rasgos de un murciélago (que tengo alas, que vuelo al atardecer, que como frutas e insectos, que tengo muy mala visión, que duermo colgado de los pies, etc.) pero esto no sería equivalente a saber cómo es ser murciélago *para un murciélago*, sino cómo sería *para mí* ser un murciélago. De nuevo el problema es que nuestras capacidades imaginativas están restringidas por nuestras experiencias y recursos mentales, los cuales son insuficientes para imaginarnos, y más aún, para saber cómo es la fenomenología de un murciélago; esto es, *cómo es ser un murciélago para un murciélago*, solo lo puede experimentar un murciélago. En particular, sin acceso a la experiencia fenoménica ecolocalizadora, no tenemos acceso al carácter subjetivo de esa experiencia. Podemos ofrecer todo tipo de descripciones (por ejemplo podemos describirlas como experiencias tridimensionales auditivas o análogamente extender el valor de la experiencia que tenemos del aparato sonar de un submarino) pero esto no bastará para saber cómo es tener esas experiencias, y por lo tanto, tampoco para saber cómo es ser un murciélago. La fenomenología de una experiencia ecolocalizadora, y por ello, la experiencia ecolocalizadora misma y el resto del cómo es ser un murciélago para un murciélago es inaccesible para nosotros. Incluso, de acuerdo con las cuatro características básicas de una experiencia subjetiva, si un murciélago adquiriera la capacidad de hablar, por muy creativo y elocuente que sea y por muy cooperativos e imaginativos que resultáramos nosotros, él no podría transmitirnos su propia experiencia de ser murciélago, su experiencia estaría en las mismas condiciones de infabilidad que la nuestra.

Esta es la *perspectiva de la primera persona*, un organismo es consciente en tanto que posee un punto de vista particular sobre el cual los objetos del mundo exterior generan experiencias determinadas (experiencias diferentes para organismos diferentes). Las rocas y las plantas estarían excluidas de la conciencia, ¡no hay nada que sea ser como una roca! Presumiblemente no se siente nada. De esta manera, insistimos, lo esencial de una experiencia subjetiva es la manera en que se siente o se experimenta, dado el punto de vista particular del organismo que lo sienta (*e. g.* ¿qué quedaría de una experiencia dolorosa si le restamos la cualidad de ser sentida como dolorosa?). Si existe algún organismo que no

tenga experiencias conscientes, entonces, según se sigue de la idea de Nagel, no hay nada que sea ser como ese organismo, ni para ese organismo ni para nada ni nadie más.

Lo que muestra esta manera de ver las cosas es simplemente que hay experiencias que son esencialmente subjetivas, es decir, en palabras de Thomas Nagel, que dependen del punto de vista del organismo que las tiene; y además que cualquier experiencia subjetiva es real en tanto exista un ser consciente con una particular *perspectiva en primera persona* para sostenerlas, en tanto que es consciente de ellas.

I.3. Los datos en primera persona: contenedores de la experiencia subjetiva

En programación computacional un dato es una entidad individual carente en sí mismo de contenido semántico; por sí solo no significa nada, adquiere significado según el tipo de *ítems* que lo integren. Un dato puede contener información sobre un número, un signo, una letra o cualquier símbolo que represente un suceso. Si obtuviéramos datos llenos de información de la experiencia subjetiva de un sujeto consciente, por ejemplo un dato sobre su experiencia del color de los tomates o un dato acerca de una creencia o un deseo, es decir datos singulares sobre su experiencia subjetiva, sobre lo que constituye su *manera de ser* en general, obtendríamos lo que se conoce como *datos en primera persona*.

Permítasenos en este punto transportar la reflexión a un asunto de gran importancia para la filosofía de la mente y la conciencia: la representación, una pieza clave para la introducción de otras experiencias en la categoría de subjetivas. Un estado mental, como una percepción, una sensación, creencia o deseo es un estado de la mente que representa al mundo de alguna manera. Cuando decimos que las creencias, deseos, esperanza y demás representan el mundo, nos referimos a que todas estas experiencias se dirigen a algo. Todos los casos en que se cree, se anhela o se desea, presentan una cualidad de direccionalidad: se cree en algo, se anhela algo o se desea algo. No tiene sentido suponer que alguien simplemente anhela, sin anhelar algo. Aquello que se crea o se anhele es lo que es representado por su anhelo. A esta cualidad de los estados mentales representacionales se le ha adoptado de distintas maneras, por ejemplo Bertrand Russell les llamó “actitudes proposicionales”. Si un estado mental tiene la forma “A” “Ψ” que S, en donde “A” es la persona que se halla en el estado mental, “Ψ” represente la actitud y “S” ocupe el lugar de

la situación representada, entonces ese estado mental es una actitud proposicional. Otro fragmento de terminología que ha sido ampliamente utilizado es el término “contenido” usado donde Russell usó “proposición”. De acuerdo con esta terminología cuando anhelo la erradicación de la violencia generada por el crimen organizado, el *contenido* de mi anhelo es *la erradicación de la violencia generada por el crimen organizado*.

En filosofía existe un término que expresa esta cualidad representacional de los estados mentales de manera precisa y técnica: la intencionalidad, y los estados mentales que exhiben intencionalidad, son llamados “estados intencionales”.⁷ Expliquemos un poco de qué se trata este asunto, comenzando con la idea más sencilla posible. La intencionalidad, con ‘c’ es la direccionalidad acerca de algo, es común encontrarse con el término “acercidad” como referente de la expresión “acerca de” en los textos filosóficos que tratan el tema. Decimos entonces que los estados mentales intencionales tienen “acercidad” porque hablan acerca de cosas. Franz Brentano por ejemplo afirma que la esencia de lo que distingue fenómenos mentales de fenómenos físicos, es que mientras los primeros exhiben esta direccionalidad, ningún fenómeno físico lo hace.⁸ Brentano sostuvo que los fenómenos mentales son caracterizados:

Por lo que los escolásticos de la Edad Media llamaron intencionalidad [...] inexistencia del objeto, y que nosotros, aunque con expresiones no del todo sin ambigüedad, llamaríamos relación con un contenido, dirección hacia un objeto, (que no está aquí para ser entendido como realidad) u objetividad inmanente. (Crane, 1995/2008, pág. 64)

Los rebuscados términos de “inexistencia intencional”, “relación con un contenido” y “objetividad inmanente” son todos ellos diferentes maneras de expresar la misma idea de

⁷ Evitemos pasar por alto la diferencia entre “Intencionalidad” e “Intensionalidad”. Recordemos que un contexto lingüístico o lógico es intensional (con s) cuando no es extensional. Tim Crane lo explica de la siguiente manera: “Un contexto extensional es uno acerca del cual son verdaderos los siguientes principios: (A) El principio de intersustituibilidad de las expresiones correferentes y (B) El principio de generalización existencial. El primero nos dice que si un objeto tiene dos nombres, N y M, y se dice algo verdadero acerca de ello usando M, no puede convertirse esta verdad en falsedad sustituyendo M por N. El segundo principio nos dice que podemos inferir que algo existe a partir de un enunciado hecho acerca de ello. Por ejemplo, a partir del enunciado ‘Orwell escribió *Animal Farm*’ podemos inferir que ‘Existe alguien que escribió *Animal Farm*’”. (Crane, 1995/2008, pág. 66) Si fracasa alguno de estos dos principios podemos decir que el contexto lingüístico es intensional.

⁸ A esta afirmación de que la intencionalidad es la “marca de lo mental” se le llama en ocasiones *Tesis de Brentano*.

que los fenómenos mentales implican representación o presentación del mundo. Tim Crane (Crane, *La mente mecánica: Introducción filosófica a mentes, máquinas y representación mental*, 1995/2008, págs. 64, 65) explica que “Inexistencia” se toma para expresar la idea de que el objeto de un pensamiento, es decir “de lo que se trata un pensamiento” existe *en* el acto de pensar, en sí mismo. Esto no es decir que cuando pienso en mi perro, haya un perro “en” mi mente. Más bien, es precisamente la idea de que mi perro es intrínseco a mi pensamiento, en el sentido de que lo que lo hace el pensamiento que es, es el hecho de que tenga a mi perro como objeto.

Franz Brentano afirmó que todos y nada más que los fenómenos mentales exhiben intencionalidad. Afirmar esto es afirmar que 1) todos los estados mentales exhiben intencionalidad y 2) únicamente los estados mentales exhiben intencionalidad. O para una mayor precisión ser un estado mental es suficiente para tener intencionalidad, y ser un estado mental es necesario para tener intencionalidad. Es tentador pensar que es verdad sin más la afirmación de que es suficiente la mentalidad para la intencionalidad, pero analicemos una problemática asociada a ello. El dolor es un estado mental de género del cual sólo un ser consciente puede ser partícipe. Como es consciente puede, por definición, echar una mirada dentro, es decir, ejercer introspección y descubrir que tal dolor no es precisamente un estado con direccionalidad hacia algo; que no es la clase de cosas que son por ejemplo los pensamientos ocurrientes. Los dolores son sólo sensaciones, no se tratan de nada ni hablan “acerca de” nada. El asunto es que aunque no diríamos que un dolor trata “acerca” de algo, exhiben un modo de representar el mundo en el sentido de que genera determinada experiencia subjetiva; se siente en mi brazo izquierdo; existe un dato lleno con el ítem de mi dolor. Esta localización es presumiblemente el factor que la hace intencional.

La idea de que la mentalidad es necesaria para la intencionalidad conlleva a un trabajo más arduo. Para sostener que las mentes no son las únicas que exhiben intencionalidad, basta con mostrar algo que la exhiba que no sea una mente; y podemos encontrar más de una instancia de este hecho. Una de ellas es el caso de los libros, cualquier libro está lleno de enunciados con sentido, que portan en sí mismos una trama de intencionalidad, representan cosas y hablan “acerca” de cosas. Pero no diríamos que un libro tiene mente. Los libros no tienen intencionalidad intrínseca, únicamente la tienen a causa de que se interpretan por un lector. Es la diferencia entre intencionalidad “original” e

intencionalidad “derivada”; la intencionalidad presente en un libro es derivada de los pensamientos de aquellos que escribirán y leerán el libro. La de la mente es original ya que no depende de la intencionalidad de nada más. El asunto sobre qué clase de cosas puede tener intencionalidad original, es tema que abordaremos en el segundo capítulo. Por ahora es suficiente con expresar que los contenidos representacionales bien pueden jugar el papel de los datos en primera persona capaces de actuar como piezas de información relevantes para una explicación cabal de la experiencia consciente.

I.4. Irreductibilidad de los datos en primera persona

Supongamos lo siguiente, el universo está compuesto de un número finito de elementos químicos interactuando de manera física unos con otros; el resultado de ello es la existencia de galaxias, constelaciones, estrellas, planetas, nebulosas, quásares y cualquier cantidad de cuerpos que “nacen” y “mueren” a lo largo del tiempo. La naturaleza de cualquiera de estos cuerpos puede explicarse dentro de los límites de las ciencias físicas conocidas; podemos estudiar su temperatura, su peso, su masa y su longitud, sus interacciones electromagnéticas, gravitacionales, nucleares y termodinámicas. Sabemos adicionalmente que en algún lugar de todo ese aparato de cuerpos celestes existe la vida. Supongamos también que la vida está constituida por esos mismos elementos químicos con sus interacciones físicas, pero que la manera en que están organizados en algunos de los miembros ha hecho emerger otras características en sus cuerpos físicos. En el lugar donde los elementos químicos interactuando físicamente en los cuerpos celestes, producen luz, calor, fuerza de gravedad y demás, en los organismos vivos produce crecimiento, movimiento, reproducción y pensamiento. Todo esto constituye la concepción de la realidad completa, tanto los elementos físicos y químicos y sus interacciones, como la vida con sus elementos físicos y sus elementos supuestamente no físicos, es decir, sus cuerpos y sus mentes. La pregunta es, ¿Tiene la mente en sí misma un carácter objetivo? ¿Los elementos mentales subjetivos son susceptibles de explicarse con la misma objetividad física con que la ciencia explica, por ejemplo, la naturaleza de la temperatura, la química del agua o el fenómeno genético?

¿Cuál es la objetividad física de la que hablamos? Es una creencia muy difundida la de que los hecho más básico del universo son hechos físicos y que todos los demás hechos dependen de ellos. Jaegwon Kim (Kim, 2001) llama *superveniencia* a esta dependencia. La superveniencia formaliza la idea de que un conjunto de hechos puede determinar por completo otro conjunto de hechos. En general la superveniencia es una relación entre dos conjuntos de propiedades: propiedades B –intuitivamente, las propiedades de alto nivel- y propiedades A -las propiedades más básicas de bajo nivel-. Según Kim la superveniencia de una conjunto A de propiedades en otro conjunto B puede explicarse así: es necesario que para cualquier propiedad F en A, si cualquier objeto x tiene F, entonces existe una propiedad G en B tal que x tiene G y por su parte necesariamente cualquier objeto que tiene G tiene F. Podemos decir que F superviene en G, es decir que G es la base para F. Por lo tanto x teniendo F superviene en x teniendo m(F) y y teniendo G superviene en y teniendo m(G), donde m(F) y m(G) son micropropiedades relativas a F y G. Si agregamos que hay una conexión causal adecuada entre esas micropropiedades, este razonamiento constituiría el esquema general para la reducción de una relación macrocausal.

Pensemos la superveniencia de la siguiente manera: en la explicación del fenómeno genético, por ejemplo, lo que se necesitó explicar fue la función objetiva de las características hereditarias transmitidas a través de la reproducción. En 1962 Watson y Crick aislaron un mecanismo microscópico que potencialmente pudo efectuar esta función: la réplica de cadenas de doble hélice en las moléculas de ADN. Esto les permitió entender cómo las moléculas de ADN efectúan la transmisión de caracteres hereditarios de célula a célula, y así, los fenómenos genéticos fueron explicados gradualmente. El resultado es un tipo de explicación reductiva, es decir, hemos explicado un fenómeno de alto nivel (macroscópico), el fenómeno de la transmisión de características hereditarias, en términos de un proceso de bajo nivel (microscópico) de biología molecular. Sustituyamos las características hereditarias por la experiencia subjetiva y la réplica de las cadenas de doble hélice por algún proceso neurofisiológico y lo que obtendríamos sería una explicación reductiva de la mente; es razonable esperar que este tipo de modelo funcione en la explicación de los fenómenos concernientes a la experiencia subjetiva. El problema es que los datos representacionales obtenidos de la experiencia subjetiva parecen no ser datos sobre funcionamiento objetivo. La razón de que el modelo de la superveniencia se enfrente

con dificultades a la hora de trasladarse a la explicación de la experiencia subjetiva es la dependencia causal entre los sucesos mentales (experiencia subjetiva) y los hechos físicos. Donald Davidson (Davidson, 1970/1981) afirma que existen al menos tres principios involucrados en este proceso.

El primero de ellos es el principio de interacción causal: algunos sucesos mentales (algunas experiencias subjetivas) interactúan causalmente con algunos sucesos físicos. Consideremos el ejemplo del semáforo en rojo. Alguien que conduce un auto al mirar en el semáforo la luz roja que indica alto, desea detenerse y pisa el freno. Formalmente, el hecho de que el automóvil se detenga, se debe en primera instancia a la convención de que la luz roja en un semáforo indica alto, en segundo lugar a que el sujeto percibe el color de la luz, en tercero que esto genera el deseo de obedecer la convención y finalmente la acción de pisar el freno. Si el auto se detuvo entonces varios sucesos mentales tales como percepciones, señalamientos, cálculos, juicios, decisiones y creencias desempeñan un papel causal sobre el hecho de que el auto se detuvo. Tenemos entonces un ejemplo en el que la percepción en combinación con otras experiencias subjetivas está conectada causalmente con un evento físico.

El segundo principio es el del carácter nomológico de la causalidad, nos dice que siempre que dos sucesos están relacionados como causa y efecto, existen leyes deterministas estrictas que rigen esa relación. Si es correcto que mirar el color rojo y saber que indica alto está conectado causalmente con el hecho de detenerse mientras conduces un auto y mirar el color rojo del semáforo, entonces debe haber una ley rigiendo esa conexión; si no, esto ocurriría unas veces y otras no.

Pero, el tercer principio afirma que no hay leyes deterministas estrictas sobre las cuales los sucesos mentales puedan predecirse y explicarse. Este principio se conoce como la *anomalía de lo mental*. Se puede pensar que estos tres principios son contradictorios entre sí; los dos primeros principios, el de la interacción causal y el del carácter nomológico de la causalidad, juntos, implican que al menos algunas experiencias subjetivas pueden predecirse y explicarse con base en leyes estrictas, o en términos reductivos como el fenómeno genético; en tanto que el principio de la anomalía de lo mental niega en absoluto la probabilidad.

Con base en estos tres principios Davidson argumenta que: si un suceso mental m causa un suceso físico f y, como no existen leyes psicofísicas, no puede haber un principio que apoye esta relación causal particular, entonces, dado el principio del carácter nomológico de la causalidad, la ley que la apoya deberá ser física y, por lo tanto el suceso mental que cae bajo una ley física deberá también ser un suceso físico. De acuerdo con Thomas Nagel (Nagel, Una visión de ningún lugar, 1986/1996) el problema de una explicación de la mente en términos de una reducción física objetiva, surgió debido a que algunas características de la vida mental presentan obstáculos para una completa definición de la realidad objetiva. Ya que en primera instancia, al concebir el mundo, éste se presenta ante nosotros por medio de la percepción sensorial. Nuestras percepciones sensoriales son causadas por la acción que ejercen los objetos físicos sobre nosotros, son efecto de la estimulación de nuestros cuerpos que son también objetos físicos. Las propiedades físicas de los cuerpos que causan en nosotros percepciones al pasar por nuestros cuerpos, también causan efectos diversos en otros objetos del mundo físico e incluso es posible que existan sin causar ninguna percepción en absoluto; su naturaleza en sí misma debe ser independiente y separable de su apariencia perceptual, es decir, separable de cualquier efecto perceptual que pueda causar en cualquier mente que los perciba (incluida la propia mente de Dios, si es verdad que existe). Tratamos entonces de formar una concepción de la realidad objetiva, al margen de las experiencias subjetivas que pueda causar en nosotros o en cualquier clase de organismo con características mentales.

Para Nagel adoptar una estrategia como esta implica no sólo dejar de pensar en el mundo físico desde nuestro particular punto de vista, sino dejar de pensar en él desde un punto de vista perceptual humano en general, no pensar en cómo se ve, cómo se siente, cómo suena, cómo huele o a qué sabe, qué creemos o deseamos de él. Implica dejar de representar el mundo por medio de las propiedades secundarias de los objetos físicos, para abrirle paso al estudio estructural de las cualidades primarias que subyacen, como forma, tamaño, peso y movimiento. El mundo descrito por la concepción objetiva carecería, entonces, de rasgos distintivos, ya que aunque las cosas que se hallan en él poseen propiedades, ninguna de estas propiedades son aspectos perceptuales. Todos ellos pertenecen a la mente, es decir, se revelan únicamente en la medida en que hay un ser consciente con su particular punto de vista para percibirlos. De manera que el mundo físico

tal como se supone que es en sí, no contiene puntos de vista ni nada que pueda ser aprehendido mediante un punto de vista particular. “Lo que contiene es susceptible de ser aprehendido por una conciencia racional que obtenga su información por medio de cualquier punto de vista perceptual desde el cual acceda al mundo” (Nagel, Una visión de ningún lugar, 1986/1996, pág. 26)

La concepción objetiva de la realidad es muy fructífera en distintos campos de la ciencia; después de todo hemos logrado explicar reductivamente nociones como el crecimiento y la auto-reparación celular, la deriva continental, el metabolismo, etc. No obstante dentro de la tradición filosófica de la mente se encuentra con dificultades si se propone como método para buscar una comprensión completa de la realidad mental, ya que parece ignorar algunos rasgos sobre la percepción, la sensación y la reflexión presente en la concepción de la experiencia subjetiva, e incluso bajo ciertas circunstancias los niega, por parecerle totalmente irrelevantes. Aunque por cierto hemos intentado mostrar que también pertenece a la realidad. El problema central no es si debe o no admitirse a la experiencia subjetiva dentro de la concepción de la realidad completa, sino que al admitirla, representa el más claro desafío para la idea de que la objetividad física representa la forma general de la realidad. Nagel no pretende abandonar la idea de la objetividad, pero sí proponer que esa visión no agota el estudio de los fenómenos de la conciencia. En lo que, hasta ahora, coincidimos con él.

Si lo que hemos dicho es cierto, y además la conciencia está compuesta de elementos físicos y de experiencia subjetiva asociada a esos elementos físicos, no podemos ignorar o negar la existencia de la experiencia subjetiva en virtud de alcanzar una noción objetiva de la realidad física. Podríamos concentrarnos en el estudio de las cualidades primarias de los objetos, pero parece que hacerlo no nos proporcionaría una visión unificada del mundo. Por lo tanto debemos inspeccionar qué posibilidades existen de alcanzar esa visión admitiendo dentro de ella los rasgos distintivos de la experiencia subjetiva.

Desde una perspectiva en primera persona no parece ser posible acceder por completo a la experiencia subjetiva de otro individuo. La idea es entonces, desarrollar una manera de evitar esa incapacidad que no sea el de la imaginación (ya hemos visto en la sección 1.II que imaginar la experiencia subjetiva de otro no es suficiente). Por lo tanto es necesario un método objetivo e imparcial, debemos poder pensar en nuestras propias experiencias

subjetivas desde fuera, en términos mentales no en términos físicos; es necesario un método objetivo de la mente que no abandone los conceptos básicos de lo mental y que no reduzca lo mental a lo físico. Nagel es optimista al respecto aunque con sus reservas:

Me parece que podemos incluirnos a nosotros mismos, con experiencias y todo, en un mundo concebible desde un punto de vista que no sea específicamente humano, y pienso también que podemos hacer esto sin reducir lo mental a lo físico. Por otra parte, creo que cualquier concepción de esta índole será necesariamente incompleta. Y esto significa que la búsqueda de una concepción objetiva de la realidad tropieza con límites que no son meramente prácticos, límites no superables para cualquier inteligencia puramente objetiva, no importa cuán poderosa sea. [...] lo anterior no es motivo de alarma filosófica, pues no hay razón para suponer que el mundo tal como es en sí, deba ser comprensible de manera objetiva, ni siquiera en el más amplio sentido. Algunas cosas sólo pueden entenderse desde adentro [...] Es natural querer llevar nuestras capacidades de comprensión objetiva, imparcial, a alinearse con la realidad tanto como se pueda, pero no debemos sorprendernos si la objetividad resulta en esencia incompleta. (Nagel, Una visión de ningún lugar, 1986/1996, pág. 30)

Es necesario pensar en nuestras mentes como instancias de algo general; de la misma manera en que estamos dispuestos a aceptar determinados hechos físicos como instancias de un sistema general, por ejemplo, cómo el calor o la frescura de un cuerpo están relacionados con la energía interna de un sistema termodinámico. En este caso, así como en el caso de la mente, la perspectiva humana no tiene importancia central (un extraterrestre interesado en estudiar las leyes de la termodinámica que han desarrollado los terrícolas, se encontraría con un sistema que no necesita de las nociones perceptuales humanas para comprenderlas). Debemos pensar entonces en la mente, lo mismo que en la materia, como un aspecto general del mundo.

La lección es que si sostenemos que las ciencias físicas pueden, en principio, explicar la totalidad de la realidad objetiva, y si aceptamos la idea de que los estados mentales son irreductiblemente mentales en virtud de una subjetividad ineliminable y no pretendemos, por lo tanto, que lo mental pueda representarse como un aspecto de la realidad objetiva, entonces sería necesario concederle a lo mental una realidad completa y autosuficiente, y así, concebirlo como un aspecto diferente del mundo natural. Si la conciencia está integrada por estos dos aspectos entonces ¿Es posible construir una ciencia adaptada a la naturaleza de los dos aspectos que los integre sistemáticamente en virtud de alcanzar una explicación completa? ¿Podemos obtener datos de ambos aspectos y construir un marco científico que

integre la relación que existe entre ellos para explicar la conciencia? La respuesta tentativa es que sí, aun cuando la naturaleza de alguno de los dos elementos no sea revelada por completo, al menos podemos esperar que exista una correlación entre ambos, de manera que podamos identificar su conexión al alterar alguno de los polos y esperar que el otro se modifique también.

I.5 Posturas filosóficas que apoyan la existencia de la subjetividad de los estados mentales: La inteligencia consciente reside en algo no-físico

La postura que hemos descrito hasta ahora, contiene elementos de lo que la tradición fisicista y materialista identifican como Dualismo o Mentalismo. Cuya idea central es que la inteligencia consciente existe como un elemento que escapa a cualquier descripción científica objetiva de la naturaleza, porque posee una ontología y una epistemología diferente de la de los fenómenos físicos, así que debe por consecuencia estar separada de ellos, como una clase diferente de entidad. A continuación describiremos en términos generales las posturas de las que hemos echado mano para construir este capítulo.

Dualismo

El dualismo en filosofía de la mente es la postura filosófica según la cual, la inteligencia consciente reside en (o es totalmente) algo con características no físicas. Esta postura filosófica contrasta el concepto “mente” con el concepto “cuerpo” como los dos posibles elementos encargados de sostener la conciencia. A lo largo de la historia de la filosofía, diferentes aspectos de la mente han generado una sutil variedad de dualismos. En lo que sigue vamos a caracterizar esta variedad, enfatizando que aun cuando existen distintos tipos de dualismos, todos ellos comparten la idea de que la conciencia o incluso los seres capaces de ser conscientes, son fenómenos no físicos. Una manera natural de comenzar el análisis es pensando qué tipo de elementos alguien asume como preponderantemente importantes para ser dualista. Las categorías más comunes son las sustancias y las propiedades, que dan como resultado, según se elijan, un dualismo sustancial o un dualismo de propiedades. Pero existe una tercera variedad, conocida como dualismo de predicados. Esta variedad de dualismo es una versión menos fuerte que las

otras dos, en el sentido de que es la que menos exige, por tal motivo comenzaremos con ella.

Dualismo de predicados

Alguien es un dualista de predicados si supone dos cosas: la primera es que los predicados de la experiencia subjetiva son esenciales para una descripción completa del mundo; y la segunda si supone que los predicados de la experiencia subjetiva no son reducibles a predicados pertenecientes a estados con objetividad física. Para que se presente una reducción auténtica de un predicado sobre un estado mental a un predicado sobre un estado físico, tendrían que existir leyes precisas que correlacionaran tipos de estados mentales con experiencia subjetiva, con tipos de estados físicos, de tal manera que el uso del predicado del estado mental contuviera toda la información del estado físico y viceversa; lo cual parece que requeriría una ciencia autónoma completamente nueva. Actualmente en muchas de las ciencias exactas existen muchos ejemplos de lo que se considera una reducción exitosa, uno de ellos es el caso del agua; algo es agua si y sólo si es H₂O. Si se reemplaza la palabra “agua” por el concepto preciso “H₂O” tenemos toda la seguridad de que estamos manteniendo y transmitiendo exactamente la misma información. El dualismo de predicados es sostenido por aquellos que piensan que la naturaleza cualitativamente subjetiva de la mente no es solamente otra manera de categorizar estados cerebrales o de la conducta, sino que se trata de un auténtico fenómeno emergente.

Dualismo de propiedades

El dualista de predicados considera que hay dos tipos de enunciados en nuestro lenguaje que dicen algo sobre la naturaleza de conciencia. El dualista de propiedades es aquel que afirma que estos predicados existen porque existen también las propiedades de las que esos predicados hablan, como clases esencialmente diferentes: las propiedades físicas, y las propiedades no físicas (o mentales). La idea fuerte de un dualista de propiedades es que el cerebro posee un conjunto de propiedades que no posee ningún otro cuerpo físico, por ejemplo, ser el agente de la sensación de dolor, de la percepción de los

colores, de pensar P, de desear Q, etc. Y asume que estos fenómenos son las propiedades características de la inteligencia consciente. El dualista de esta clase sostiene que estas características de la conciencia son no físicas porque no podrán ser reducidas nunca a los conceptos de las ciencias físicas conocidas, y por lo tanto no podrán ser explicadas por éstas.

Dualismo sustancial

En la noción de dualismo sustancial intervienen dos nociones importantes: la de sustancia y la de dualismo de tal sustancia. Una sustancia está caracterizada por sus propiedades, pero de acuerdo con quienes sostienen una postura como esta, una sustancia no es únicamente el conjunto de propiedades que posee, sino que de hecho, es *lo* que las posee; una sustancia inmaterial que posee estados con propiedades inmateriales. Esto quiere decir que una mente es un paquete individual de sustancia no física, con independencia de cualquier objeto físico al que pudiera estar unido. Uno de los primeros filósofos en elaborar una teoría sobre la mente apoyada en esta variedad de dualismo fue René Descartes (1596 - 1650). Para Descartes el mundo estaba dividido en dos tipos básicos de sustancias: la primera era la materia común, cuya característica principal es la de ser una sustancia extensa, es decir cualquier cosa que ocupa un lugar en el espacio, cuantificable mediante su longitud, peso, volumen y demás; y la segunda era la razón consciente del hombre, que en su opinión era el único aspecto de la realidad que no podía ser explicado en términos de la mecánica de la materia. Por ese motivo postuló la existencia de una sustancia radicalmente diferente e independiente de la sustancia física, que no posee ninguna posición espacial y cuya característica principal es la actividad de pensar, a la que denominó *res cogitans* y en la que según él residía el verdadero *sí mismo* de cualquier sujeto.

Dualismo popular

El dualismo popular es aquella postura que defiende la existencia de algún tipo de sustancia inmaterial cuya constitución interna es totalmente diferente de la materia física pero que aun así, puede presentar plenamente algunas propiedades físicas. Esta postura

lleva el epíteto de “popular” dado que es común encontrar esta opinión en algunas de las religiones más populares, cuya idiosincrasia se basa en la procuración de la vida del mundo futuro. En particular, la creencia generalizada es que el cuerpo físico posee un alma dentro de él y que es gracias a ella que cada persona posee características que lo hacen único e individual. Este enfoque resulta muy atractivo para muchos por una razón en especial, y es que por lo menos contempla la posibilidad de que el alma sobreviva a la muerte del cuerpo. Aun cuando se enfrenta a la clara resistencia de que desear que el dualismo estuviera en lo cierto dada la esperanza de la supervivencia del alma, no es una buena razón para creer que lo está.

Epifenomenismo

Es una rama del dualismo de propiedades. Sostiene que los fenómenos mentales no forman parte de los fenómenos físicos del cerebro (objeto físico que determina nuestras acciones y conductas) sino que están por encima de ellos. La creencia general es que se presentan cuando el desarrollo del cerebro supera un determinado nivel de complejidad. El epifenomenista sostiene que aun cuando la causa de que se produzcan fenómenos mentales sean actividades del cerebro, dichos fenómenos no tienen efectos causales en el mundo físico. Como consecuencia, la creencia universal de que nuestras acciones están determinadas por nuestros deseos, sería falsa; nuestras acciones estarían determinadas por hechos físicos del cerebro que también son la causa de los epifenómenos que denominamos, por ejemplo, deseos, creencias, amor, cosquillas etc. Hay una conjunción constante entre volición y acciones pero es una simple ilusión que la primera sea la causa de las segundas. La introspección es también testimonio para explicar estos hechos. Un neurocientífico ocupado de rastrear los orígenes físicos de la conducta, no podría negar que ha tenido experiencias, creencias y deseos conectados a su conducta. Podemos conceder la realidad de las propiedades mentales, entonces, pero acoplándolas a la categoría de epifenómenos (entidades que no tienen ningún poder explicativo de la conducta en general).

Dualismo interaccionista de propiedades

Un dualista de esta categoría sostiene que las propiedades mentales efectivamente tienen efectos causales sobre el cerebro y debido a eso también sobre la conducta. De la misma forma que el epifenomenista, el dualista interaccionista asevera que las propiedades mentales son propiedades emergentes, es decir, que no aparecen de ninguna forma hasta que la materia física común haya podido organizarse evolutivamente para formar un sistema lo suficientemente complejo para sostenerlas.

CAPÍTULO II

LA OBJETIVIDAD DESEADA, MATERIALISMO Y LOS DATOS EN TERCERA PERSONA

II.1. Conductismo

El Conductismo es considerado la primera gran corriente de pensamiento materialista del siglo XX. Durante su desarrollo adquirió dos variantes principales, en distintos campos. La primera de ellas fue en psicología, conocida también como Conductismo metodológico, y la segunda en filosofía, también denominada Conductismo analítico o filosófico. Ambas variantes poseen diferencias notables entre sí, pero comparten la misma tesis central y su aparición y desarrollo tuvieron las mismas motivaciones. Principalmente el Conductismo surgió como una reacción en contra del Dualismo, durante su auge sostuvo la idea del positivismo lógico de que las circunstancias observables verifican o refutan el significado de cualquier proposición y asumió el supuesto generalizado hasta entonces, de que la mayor parte de los problemas filosóficos, o todos ellos a los largo de la historia, son el resultado de una confusión lingüística que puede disolverse con un análisis cuidadoso de su significado y del lenguaje en el que se expresan. La tesis central del Conductismo es simple y es consecuente con las tres motivaciones de las que surgió: la expresión “estados mentales”⁹ no es sino una manera de designar la conducta de las personas. La teoría evidentemente tramaba la reducción de la gama de experiencias subjetivas del mentalismo a conductas observables en tercera persona.

Tanto el Conductismo Psicológico como el Filosófico adoptaron posturas diferentes ante la misma tesis central. La más notable de ellas consiste en que por un lado el Conductismo Psicológico la consideró como una hipótesis de carácter empírico que debe ser demostrada mediante la experimentación, lo que llevó al desarrollo de teorías que destilaron del objeto de estudio las impurezas de conceptos como alma, mente o incluso conciencia, para ocuparse del estudio directo de los organismos en interacción con sus factores ambientales; de los datos que revelan su naturaleza objetiva. Por su parte los conductistas filosóficos trataron a la tesis como una verdad analítica a priori que no

⁹ Lo que aquí hemos también llamado experiencia subjetiva.

requería demostración, bajo la cual estarían organizadas las hipótesis potencialmente contrastables. La tesis central puede desglosarse en dos afirmaciones:

- Toda la variedad de experiencias subjetivas que popularmente son reconocidas y se identifican como “estados mentales”, en realidad no existen.
- La expresión “estados mentales” es una manera de designar la conducta de las personas.

Un “estado mental”, según el conductista, posee identidad con una conducta particular, ambas son una y la misma cosa. Así que la manera en que se consideran los estados mentales popularmente, es engañosa, puesto que induce a la gran mayoría de las personas a reconocer la existencia de una entidad de naturaleza confusa que determina (que está, en sentido estricto, conectada causalmente con) las conductas de las personas y no únicamente las conductas mismas.

En el contexto del Conductismo, no interesa saber cuál es la naturaleza de una experiencia subjetiva, de hecho ni siquiera importa averiguar si existen o no, porque si es posible vincularlas con ciertas constantes reconocibles y manipulables, eso bastará para determinar las regularidades de los elementos psicológicos en juego en el estudio de la conciencia. Eso ocurre de igual manera con el significado, únicamente le interesan al conductista aquellos objetos cuyos significados relacionan a las experiencias subjetivas con las variables experimentales conocibles y el comportamiento. El Conductismo es una teoría sobre cómo analizar y comprender el vocabulario que utilizamos para hablar de los estados mentales, o en específico de las experiencias subjetivas. Cuando hablamos de emociones, sensaciones o percepciones y de creencias o deseos no hablamos sobre episodios inmateriales, en su lugar se trata de articular un modo de hablar sobre modelos reales o potenciales de conducta. El estudio de los conceptos que postuló el Conductismo permitió dar cuenta de la mente sin apelar a entidades *misteriosas* inmateriales, cuya existencia fuera imposible de corroborar mediante la experimentación. Tales entidades, reconocía el conductista, no son privadas, inefables ni intrínsecas, sino de dominio público y con descripciones en términos físicos, como aquellos hechos visuales y auditivos o las conductas resultantes; o incluso la historia de eventos que hacían a una persona preferir

algo por encima de otra cosa. Esto permitió al Conductismo contar con el recurso científico de la predicción de hechos y así comenzar a ser, no sólo una teoría cualitativa, sino considerarse como parte de las ciencias empíricas cuantitativas que estudian la mente.

De acuerdo con esta tesis una explicación de la conciencia en estos términos hacía plausible diferencias en la conducta de dos individuos distintos ante un mismo estímulo ambiental; apelando al panorama de reforzamiento histórico. Si la conducta de dos individuos X y Y son diferentes ante el mismo estímulo, se debe a que han sido sometidos a lo largo de su vida a un mismo estímulo, bajo condicionamientos diferentes, un condicionamiento constante en el caso de X y un condicionamiento constante diferente en el caso de Y; de ahí la diferencia en las conductas. Planteando así el funcionamiento de la conducta de un individuo, no sólo en relación a eventos aislados basados en estímulos aislados, sino introduciendo, al mismo tiempo, el número de condicionamientos y refuerzos históricos, se desprende, lo que podría ser otra razón del éxito. Hablamos de la posibilidad de responder afirmativamente al problema de la existencia de otras mentes, y para decirlo de una vez por todas, en el contexto del Conductismo, si algo se comporta como un ser humano, entonces necesariamente eso es un ser humano. A diferencia de la razonable duda escéptica dualista que hundía en el más profundo solipsismo a toda mente con la capacidad de percibir indudable y directamente su actividad mental, pero nada más que la suya, la respuesta del Conductismo es que si hay dos conductas distintas ante el mismo estímulo, se debe a que detrás de cada conducta hay una vida individual con respuestas conductuales diferenciadas.

A diferencia de los conductistas de corte psicológico, los conductistas filosóficos no enfocaron sus esfuerzos en la hipótesis científica potencialmente comprobable a posteriori, en la que el progreso de su disciplina como una ciencia cuantitativa y práctica, se daría mediante la experimentación, sino que optaron por reconocer en la hipótesis de que los estados mentales de una persona son su conducta misma, una verdad analítica a priori que no es necesario comprobar, por el contrario, debe ser considerada como el axioma básico a partir del cual se construiría el conocimiento de la psicología. Este enfoque fue el resultado de la influencia del pensamiento del positivismo lógico del Círculo de Viena, que a partir de la tesis propuesta por Wittgenstein según la cual “La totalidad de las proposiciones verdaderas es la ciencia natural entera” (Wittgenstein, 2007, pág. 160) dictaba que el criterio

para determinar la verdad de cualquier proposición es que pudiera ser comprobada empíricamente o bien ser un juicio analítico a priori; un elemento al que pudieran reducirse las proposiciones de las ciencias particulares y así combatir al escepticismo radical. El positivismo lógico fundado en torno a las discusiones epistemológicas en defensa del empirismo, el rechazo de la metafísica y la unificación de la ciencia en sus lenguajes y métodos, hizo posible que el Conductismo desarrollara sus tesis como las conocemos.

De acuerdo con las ideas postuladas por el positivismo lógico, la unificación de las ciencias se enfrentaba con el problema de que cada una de ellas había desarrollado lenguajes particulares tomados del lenguaje natural, con la función de permitir ciclos del lenguaje exitoso pero que al ser utilizado en lenguaje científico resultaban ambiguos. Rudolf Carnap en *La construcción lógica del mundo* (Carnap, 1988), ideó una solución que consistía en proponer el uso de un lenguaje sígnico que no utilizara el mismo signo con símbolos distintos,¹⁰ ni usara de una misma forma signos que designaran de modo diferente, al que nombró lenguaje de constitución. El lenguaje de constitución tenía por objetivo sentar las bases de un sistema que unificara los objetos de todas las ciencias de acuerdo con la reductibilidad de un objeto a otro; así si todos los lenguajes de las ciencias particulares pudieran ser reducidos a un único lenguaje constitutivo, la unificación de las diversas ciencias —era sólo cuestión de tiempo. ¿Qué se requería para que los distintos lenguajes de las ciencias particulares, pudieran ser reducidos a un solo lenguaje? La respuesta de Carnap señala que si para cada objeto hay un hecho básico, entonces hay bases de existencia que son condiciones necesarias y suficientes para que ese hecho sea lo que es. Si las ciencias naturales estudian hechos básicos de objetos particulares, la proposición que expresa las condiciones necesarias y suficientes debe estar expresada mediante una verdad analítica a priori. El resultado sería que las ciencias particulares comenzarían a ambicionar la necesidad de proporcionar las condiciones necesarias y suficientes para garantizar la verdad de los hechos básicos de sus proposiciones; en pocas palabras encontrar su axioma básico en una proposición analítica a priori y basar sus experimentaciones en ella. Un ejemplo de la física que utiliza Carnap para ilustrar este proceso es más o menos el

¹⁰ Recordemos de nuevo las proposiciones 3.32 y 3.321 del *Tractatus lógico-philosophicus*: “Un signo es lo que, en un símbolo, es perceptible por los sentidos” y “Dos símbolos distintos puede por tanto tener en común el signo (escrito, sonoro, etc.); significan entonces de modo y manera distintos” (Wittgstein, 2007, pág. 136).

siguiente: el equilibrio de temperatura es uno de los objetos pertenecientes a la termodinámica, disciplina científica particular. El hecho básico es “x tiene respecto a y un equilibrio de temperatura”; saber esto es posible mediante una proposición que exprese un evento experimental “si los cuerpos x y y son puestos en contacto espacial directo, no se presentará ni un aumento ni una disminución en la temperatura”. La prueba de que esta proposición expresa las condiciones necesarias y suficientes del hecho básico es que ambas proposiciones son co-extensionales, el significado de una proposición está contenido analíticamente en el significado de la otra proposición, es decir, que si integramos la proposición “x tiene respecto a y un equilibrio de temperatura” al lenguaje de constitución definiéndolo a través de la proposición “si los cuerpos x y y son puestos en contacto espacial no se presentará ni aumento ni disminución de la temperatura”, o viceversa, el significado se mantiene. De esta manera se comienza a vislumbrar la intención del Conductismo de desarrollar, en relación con el problema de explicar la naturaleza de la conciencia, una tesis como esta. Si en ciencias como la física o la química, es posible encontrar hechos básicos expresados en proposiciones verdaderas analíticas a priori, es natural esperar que en la psicología como una ciencia que estudie cualitativa y empíricamente su objeto, ocurra lo mismo. Los conceptos teóricos postulados para llevar a cabo esta tarea son “conducta”, “estímulo”, “refuerzo” y más tarde “disposición”; además de que, en una visión global de Conductismo, la proposición “los estados mentales no son otra cosa que la conducta” hacía las veces de una verdad analítica a priori comprobable mediante la experimentación.

En sus inicios el Conductismo postuló que oraciones del tipo “Alfredo siente que nadie entiende su visión del rojo 2, por eso, al hablar con otras personas, le llama rojo”¹¹ son ejemplos de un mal uso del lenguaje porque el significado de la oración subordinada “Alfredo, al hablar con otras personas, le llama ‘rojo’ al ‘rojo 2’” depende del significado de la oración “Alfredo *siente* que nadie entiende su visión del ‘rojo 2’” que es una oración sobre una experiencia subjetiva que contiene elementos arbitrarios no referentes únicamente a descripciones de conductas, de manera que debían ser remplazadas por oraciones del tipo “Alfredo, al hablar con otras personas, llama ‘rojo’ al ‘rojo 2’”. Pero

¹¹ Oraciones como estas constituyen el mejor ejemplo para determinar la conexión causal que existe entre algunos sucesos mentales y algunos sucesos físicos.

pronto notaron que si bien las conductas de las personas no podían ser causadas por experiencias subjetivas, debía existir algo de lo cual dependieran, porque de lo contrario cualquier persona en cualquier circunstancia podría tener cualquier conducta; mucha más arbitrariedad de la que hay con la dependencia en la experiencia subjetiva. Así que todo resultó en el desarrollo del concepto “estímulo” para determinar el tipo de conducta que presentaba un individuo a determinado evento. Esta estrategia garantizaba una postura materialista, ya que las únicas entidades aceptables para el estudio de la mente se basaban en estímulos y conductas; conceptos que podían, según los conductistas, describirse en términos puramente físicos, totalmente objetivos. Lo que en este contexto constituye los datos en tercera persona del Conductismo.

Una conducta o un estímulo, a diferencia de un “estado mental”, son elementos que permiten la elaboración de una teoría psicológica comprobable empírica y objetivamente; ellas no se encuentran más allá de las posibilidades de la experimentación directa. Una teoría de la mente que asuma además de los datos objetivos de la conducta, la existencia de entidades subjetivas, carecerá de valor científico según el Conductismo.

Los elementos psicológicos, incluidas las experiencias subjetivas desglosadas en el capítulo I, poseen rasgos evidentemente cualitativos y abstractos; parece a primera vista que no pueden ser captados por un estudio objetivo; consideremos por ejemplo el “hecho básico” de un dolor dentro de la psicología tradicional, el dolor se define principalmente por la manera como es percibido por un individuo (*i.e.* por su dolorosidad) lo que sugiere que es un “hecho básico” puramente subjetivo, característica que automáticamente excluye a otro individuo de la posibilidad de sentir ese dolor. Si se considera como lo hace Carnap y el resto de los positivistas lógicos que la psicología es una ciencia empírica igual que las demás, es decir, una ciencia “sin más parentesco con la filosofía que cualquier otra ciencia natural”, entonces es razonable esperar que el proceso de constitución de sus objetos se vuelva realizable.

¿Es correcto tomar elementos como la experiencia subjetiva del dolor, como hechos básicos de la psicología en el marco de su transformación científica? Carnap, afirma al respecto que “en la psicología todavía no hay unanimidad respecto de los principios fundamentales que deben servir como guía de la investigación” (Carnap, 1988, pág. 235). La imposibilidad de la unificación de los conceptos de la psicología se debe poco a que no

se hayan encontrado condiciones necesarias y suficientes de los hechos básicos, y más a una tensión entre dos corrientes subyacentes al Conductismo. Por un lado, se adquieren los lineamientos de los positivistas herederos del pensamiento empirista de los siglos XVII y XVIII, que tienden a señalar que los únicos hechos básicos a partir de los que es posible construir teorías sobre la naturaleza de la mente son los “*sense-data*”¹², es decir, las experiencias subjetivas; y por otro lado el reconocimiento de que la mayoría de los experimentos de las ciencias particulares exactas como la física y la química, sugieren que deben ser hechos puramente objetivos aquellos que constituyan los hechos básicos. En la primera etapa de su pensamiento como positivista lógico Carnap asumió que los objetos físico eran en teoría reducibles a los mentales y viceversa, por lo que era posible en principio determinar las bases del sistema de constitución de los problemas alrededor de la mente, pero pensaba también que legitimar teorías materialista de la mente como el Conductismo, en términos de la práctica experimental objetiva, no era posible porque resulta problemático justificar adecuadamente la reproducción experimental de la ordenación de los objetos, en la manera en como los percibimos y la manera en la que son en realidad.

El Conductismo basado totalmente en las intenciones del positivismo lógico, pronto se enfrentó con críticas a problemas que no podía resolver; por ejemplo el hecho de asumir que los “hechos básicos” que debían servir para constituir los objetos de la psicología se encontraban condenados a permanecer incommunicables y privados. Además tampoco podía aportar una solución clara al problema de las otras mentes como el propio Carnap señala: “en principio no es posible observar en otra persona los ‘hechos básicos’ del problema psicofísico” (Carnap, 1988, pág. 309) lo que arroja por supuesto no sólo el problema de la

¹² En la Stanford Encyclopedia of Philosophy (puede consultarse en línea en <http://plato.stanford.edu/entries/sense-data/>) “*Sense-data*” se definen como “[L]os presuntos objetos dependientes de la mente de los que nos damos cuenta directamente y que tienen exactamente las propiedades que parecen tener. Por ejemplo, los teóricos de los *sense-data*, dicen que, tras mirar un jitomate en condiciones normales, uno forma una imagen del jitomate en su propia mente, esta imagen es roja y redonda. La imagen mental es un ejemplo de un ‘*sense-datum*’. Muchos filósofos han rechazado la noción de *sense-data*, porque creen que la percepción nos da una conciencia (*awareness*) directa de los fenómenos físicos, en lugar de simples imágenes mentales o porque creen que los fenómenos mentales implicados en la percepción no poseen las propiedades que aparecen ante nosotros (por ejemplo, podría tener una experiencia visual que representa un jitomate redondo y rojo, pero mi experiencia no ser en sí misma redonda y roja)”.

duda sobre otras mentes, sino además la duda sobre si además hay otros “hechos básicos” en las otras mentes.

La tesis esencial de Conductismo en general, y en particular en el Conductismo filosófico, parece apuntar hacia una reducción de los objetos (*i.e.* de las proposiciones) de la psicología a sus “hechos básicos” determinados por sus condiciones de posibilidad necesarias y suficientes; que son principalmente proposiciones sobre conductas visibles, mediante una reducción de los términos psicológicos a los términos del lenguaje de constitución. Ello derivó en la necesidad de desarrollar, al mismo tiempo, definiciones detalladas de los “hechos básicos” para poder traducir las proposiciones del lenguaje natural al lenguaje científico. Lo que hizo posible que se desarrollara el argumento que muestra la validez lógica del Conductismo:

- El significado de una proposición está determinado únicamente por las condiciones que verifican dicha proposición.
- Si una proposición tiene un contenido intersubjetivo (*i.e.* un contenido que puede ser compartido por varias personas), entonces las condiciones que las verifican deben ser públicamente observables.
- Sólo los datos de la conducta y los fenómenos físicos son públicamente observables.
 - Por tanto, el contenido de cualquier enunciado psicológico debe poder ser determinado a través de proposiciones que contengan únicamente datos de la conducta y descripciones de fenómenos físicos. Estas proposiciones deben ser verdaderas si el contenido del enunciado psicológico en cuestión es considerado verdadero también.

En el caso de la psicología se intentó proporcionar una definición explícita de las entidades mentales en términos de estímulos, de las experiencias subjetivas (*sense-data*) y de las conductas en términos de los estímulos. Sin embargo el programa conductista fracasó porque resultó siempre imposible, en el caso de los estímulos, justificar que una sensación de color es idéntica a la radiación de fotones ordenados de determinada manera. En el caso

de las experiencias subjetivas, además de ser ignoradas y relegadas al campo de la metafísica, no existía la posibilidad de reducir las proposiciones resultantes de su expresión y a su vez de los diversos lenguajes de las ciencias naturales a un mero lenguaje de vivencias personales. Y finalmente en el caso de las conductas, el programa conductista también fracasó en tanto que los síntomas visibles que indican una emoción como la tristeza, son confirmables y medibles en términos de respuestas fisiológicas corporales que nada tienen que ver con la conducta, que incluso se presentan sin que se presenten necesariamente respuestas conductuales asociadas.

La mayoría de los errores del Conductismo, hasta esta etapa, se debieron a que su programa fue demasiado ambicioso. Elaborar un lenguaje universal de constitución al cual pudiesen ser reducidos todos los términos de las ciencias empíricas exactas como la física y la química; e incluir en esta ambición también a disciplinas en las que presumiblemente están presentes elementos subjetivos, como la psicología, fue una tarea cuasi imposible. Por esos motivos hubo quienes, siendo partidarios de la tesis central del Conductismo, hicieron planteamientos con ambiciones mucho más modestas, y por ello, con un alcance más sólido.

Plantear el tema de la “conciencia” y la “subjetividad” dentro de ésta y las corrientes posteriores, levantaba siempre la sospecha de defender un paradigma cartesiano –el mentalismo- que presumiblemente carecía para entonces de importancia alguna; ya que para las corrientes como el Conductismo y posteriormente la Teoría de la Identidad, el Funcionalismo y el Eliminativismo, representaban únicamente problemas mal planteados, o mejor dicho, falsos problemas. Gilbert Ryle fue uno de los defensores de esta idea, en su libro *El concepto de lo mental* (Ryle, 2005), elaboró una teoría completa de la mente por medio de la reformulación del problema central de la filosofía de la mente en términos distintos. Ryle reveló que el problema surge elementalmente de una confusión en el uso de los conceptos “procesos físicos” y “procesos mentales”, que denomina peyorativamente “el dogma del fantasma en la máquina” el cual, según él, representa la Doctrina Oficial para la filosofía de la mente. De acuerdo con Ryle el error del mentalismo radica en afirmar que lo propio de la subjetividad es su irreductibilidad a criterios físicos, inaccesibles al dominio intersubjetivo u objetivo. Recordemos que Descartes planteó, en efecto, que “para ser no tenemos necesidad de extensión, de figura, de ser en algún lugar, ni de alguna cosa

semejante que se pueda atribuir al cuerpo, y manifiestamente conocemos que nosotros somos en razón sólo de que pensamos” (Descartes R. , 1644/1995, págs. 25-26).

La Doctrina Oficial procedente principalmente de Descartes sostiene que todo ser humano *tiene* un cuerpo y una mente, ambas cosas; o bien *es* un cuerpo y una mente, y esa dualidad constituye la totalidad de la naturaleza de un ser humano. Cuerpo y mente están unidos, pero después de la muerte del cuerpo, la mente puede continuar existiendo y funcionando. El cuerpo humano está en el espacio sujeto a leyes mecánicas objetivas y las cosas que le pasan pueden ser manipuladas por observadores externos. Las operaciones de la mente por el contrario no son observables y su desarrollo es privado, únicamente el individuo tiene conocimiento directo de los estados de su mente. En consecuencia cada individuo posee dos vidas paralelas, una constituida por lo que ocurre a su cuerpo (proceso físico) y otra por lo que le ocurre a su mente (proceso mental). De acuerdo con la Doctrina Oficial todo individuo puede aprehender directa e indubitablemente los eventos de su vida mental; y al mismo tiempo tiene también conocimiento directo e indubitable de ella mediante la introspección, como revisamos en el capítulo anterior. Los estados y procesos conscientes no son susceptibles de dudas, se revelan inevitablemente como una corriente de pensamientos, sentimientos, deseos, percepciones, recuerdos e imágenes. Se acostumbra expresar que aquellos eventos que le ocurren al cuerpo son “externos” y lo que pasa en la mente es “interno”. Bajo la presentación metafórica espacial existe una suposición más profunda. Se presume que hay dos tipos diferentes de existencia, la existencia física y la existencia mental; lo que tiene existencia física está compuesto de materia o es resultado de ella y/o de su interacción y lo que tiene existencia mental posee conciencia o bien es una función de ella.

Las implicaciones teóricas de la Doctrina Oficial según Ryle son que la ubicación espacial de los objetos materiales provoca que sus interacciones estén conectadas mecánicamente y regidas por leyes físicas; por el contrario los hechos mentales se mantienen aislados y no existe conexión causal directa entre lo que le ocurre a una mente y lo que le ocurre a otra. Se puede articular que la mente de una persona puede llegar a afectar la mente de otra persona únicamente a través del mundo físico. En el mismo esquema de la Doctrina Oficial, otra implicación es que si construimos lingüísticamente los

conceptos comunes referentes a facultades y operaciones mentales de los demás sin ningún problema, entonces los verbos, sustantivos y adjetivos con los que en nuestra vida cotidiana describimos el comportamiento de los demás, son entendidos como haciendo referencia a episodios que ocurren en historias secretas ajenas a nosotros o bien como significando disposiciones que hacen que tales episodios ocurran; y eso no parece ser el caso, que en la teoría sea tan sencillo afirmar o aun justificar que sea posible. Cuando decimos que Alfredo siente que nadie entiende su visión del “rojo 2” y por eso le llama “rojo”¹³ denotamos el acontecimiento de modificaciones específicas en lo que se supone que es inaccesible para nosotros, en tanto que se trata de una manifestación de algo en la corriente oculta de la conciencia de Alfredo. Suponemos que únicamente su acceso privilegiado podría producir testimonio auténtico de que ese uso lingüístico está aplicado correcta o incorrectamente. No obstante, resulta evidente que todos sabemos en general, cómo aplicar tales usos lingüísticos apropiadamente y cómo corregirlos cuando resultan confusos, así que fue necesario construir una teoría que enfatizara esta peculiaridad.

Al igual que los conductistas psicológicos y los conductistas filosóficos, Ryle tuvo la convicción de que hablar de estados mentales es totalmente erróneo. Pero su estrategia no consistió en intentar realizar una reducción del lenguaje de la psicología mentalista a un lenguaje de constitución como en el amplio programa de Carnap y los positivistas lógicos, sino desentrañar las razones por las que la Doctrina Oficial se originó y así afirmar que la dimensión ontológica que se asume con ella en realidad es irrelevante. En la opinión de Ryle la Doctrina Oficial reside no solamente en un conjunto de errores, sino en un gran error de tipo muy particular: un error categorial, el cuál presenta los hechos de la vida mental como si pertenecieran a un tipo o categoría lógica, cuando en realidad pertenecen a otra. La mente no es una entidad más allá de los síntomas visibles (i.e. estímulos y conductas) a partir de los cuáles se deduce tradicionalmente su existencia: lo que denominamos tradicionalmente “mente” es sólo la manera en la que están organizados estos síntomas visibles. Ryle propone ejemplos de lo que llama error categorial, que son fácilmente reproducibles, uno de ellos puede ser el siguiente: un chico que observa el paso de una marcha de manifestación y que después de señalársele al contingente del sindicato, a

¹³ De acuerdo con la perspectiva desde la que Ryle aborda el Conductismo el concepto “siente” deja de ser una experiencia subjetiva o un episodio fantasmal, para convertirse en una disposición de conducta.

los estudiantes y las mantas de apoyo, después de escuchar las consignas cantadas con rimas contagiosas y los performance artísticos urbanos, preguntara -¿cuándo presenciare la manifestación?- Suponiendo que es un elemento distinto de los que ya ha observado, se le mostraría su error diciéndole que al ver pasar a los contingentes, escuchar las consignas y demás, estaba presenciando la manifestación. La manifestación es la manera en la cual están organizados los contingentes, las mantas, las consignas y los performance. Con el caso de la mente ocurre algo similar, las personas que están convencidas de que existe y que se preguntan cuál es su naturaleza, cometen el mismo error que el chico del ejemplo. Esto es un error categorial.

Como señalábamos, en sus inicios el Conductismo sostenía que la mente no era otra cosa que el conjunto de conductas actuales que un sujeto realizaba públicamente. Pero en la etapa en la que Ryle redactó su obra, se mantenía en vigor el argumento anti-conductista según el cual había personas que sentían dolor aun cuando no podían o no deseaban expresarlo a través de algún comportamiento visible, para lo cual modificó la estrategia conductista. En lugar de considerar la “mente” de una persona como siendo el conjunto de conductas ejecutadas al momento, debían ser consideradas como el conjunto de conductas potenciales o *disposiciones de conducta*, que un sujeto tiene ante determinado condicionamiento o evento ambiental. Esta modificación tenía una ventaja esencial frente a la estrategia conductista del empirismo lógico: permitía dar cuenta de los estados mentales aun cuando no fuera posible observar comportamientos que revelaran la presencia del determinado episodio oculto; y se mantenía la estrategia materialista. El paralelismo científico de esta estrategia es la fragilidad de una lámina de vidrio, la cual puede ser determinada sin necesidad de comprobarla rompiéndola; la fragilidad del vidrio no consiste en el hecho de que en un momento dado podamos hacerlo pedazos, puede ser frágil sin que se haya roto. Esta es una definición operacional en la que el término “frágil” se define en términos de ciertas operaciones o pruebas que podrían revelar si el término verdaderamente se aplica o no en el caso que se ha de examinar. En palabras directas de Ryle:

[P]oseer una propiedad disposicional no consiste en encontrarse en un estado particular o experimentar determinado cambio. Es ser susceptible de encontrarse en un estado particular o de experimentar un cambio cuando se realiza determinada condición. Lo mismo puede afirmarse de disposiciones específicamente humanas como son las cualidades de carácter. Ser un fumador no implica que en éste o aquel instante

esté fumando, sino que soy propenso a fumar cuando no estoy comiendo, durmiendo, leyendo, atendiendo un funeral, o cuando ha transcurrido algún tiempo después del último cigarrillo (Ryle, 2005, pág. 42).

Comparada con la estrategia delineada por el Conductismo anterior, la estrategia de Ryle resulta más inclusiva y sólida. Según la anterior versión el hecho de que Alfredo sintiera que nadie entiende su visión del “rojo 2”, sólo puede ser corroborada en el momento en el que Alfredo le llama “rojo” y emite una mueca de resignación. En cambio bajo el criterio de verdad de la afirmación de que Alfredo siente algo, con este nuevo enfoque consiste en la verificación de una serie de oraciones condicionales del tipo 1) “Si se le preguntara a Alfredo cómo prefiere hablar de su visión especial, él respondería que de la manera convencional” 2) “Si Alfredo tuviera que enseñar a alguien a mezclar los colores básicos para hacer violeta, llamaría ‘azul’ al azul y ‘rojo’ al especial rojo 2” 3) “Si Alfredo necesitara indicar el color de los tomates maduros, les llamaría rojos y no rojos 2”, etc. Para Ryle es necesario que en cada mundo posible en donde Alfredo tenga una visión especial del rojo y sienta que nadie entenderá su visión especial, le llame rojo al rojo 2 toda vez que de hecho sienta que no le entenderán. Por lo tanto, no tiene sentido preocuparse por la relación entre la mente y el cuerpo. En la medida en que Ryle considera que las disposiciones son el material que da cuenta de la explicación de los estados mentales que prescinde del lenguaje mentalista, su visión frente a la de los conductistas metodológicos y la de los positivistas lógicos resulta más sólida. Sin embargo no lo suficientemente sólida para representar una visión global y definitiva de los problemas asociados con la conciencia. Recordemos aquí que nuestro interés es reconocer que para construir una ciencia de la conciencia sólida, es necesario considerar la verdad y existencia de ambos aspectos, tanto el proceso físico, como la experiencia subjetiva asociada a ello. Desafortunadamente el Conductismo depurado de Ryle también enfrentó problemas, y mucho de ellos se deben a la manera de considerar o no, dentro de su marco teórico a las experiencias subjetivas. Lo que sin duda también provocó el desarrollo de las teorías subsecuentes.

El primer problema consiste en el hecho de que ignoró e incluso negó el aspecto subjetivo de nuestros estados mentales, como ya lo mencionábamos, tener un dolor no parece ser algo que nos lleve únicamente a lamentarnos, tomar un analgésico e ir al médico;

un dolor posee una pieza de información subjetiva e intrínseca que se pone de manifiesto en la introspección. Sin embargo un lenguaje privado para hablar de ello es absolutamente imposible. De acuerdo con el argumento del lenguaje privado de Wittgenstein si intentamos otorgarle significado a un término “A” referente a un dolor, únicamente asociándolo con una sensación que se experimenta en determinado momento. En un momento posterior en el que se experimente de nuevo tal sensación, se puede afirmar que se trata de otro “A”, pero es imposible determinar que se ha usado el término correctamente en esta segunda ocasión porque quizá no se recuerde bien la primera sensación, o desenfadadamente se perciba una semejanza entre una y otra sensación, aunque *en realidad* se trate de dos sensaciones distintas; sin una conexión de sentido con otros fenómenos, por ejemplo causas y/o efectos, no habrá ningún modo de distinguir entre el uso correcto de “A” y su uso incorrecto. Este es el problema semántico del Conductismo frente a la explicación de la conciencia.

El segundo problema apareció cuando los conductistas intentaron especificar con el máximo detalle la disposición presuntamente constitutiva de todo estado mental. La lista de condicionales necesaria para un análisis adecuado de “Alfredo siente que nadie entiende su visión del rojo 2” no sólo es indefinidamente larga, tiende a ser infinita. Incluso cada uno de los condicionales asociados a la definición es dudoso en sí mismo. Suponiendo que Alfredo siente auténticamente que nadie lo entiende, el condicional 1) antes mencionado sería verdadero sólo si Alfredo *desea* llamarle rojo al rojo 2; el condicional 2) sólo sería verdadero si Alfredo *quisiera* mantener el orden y la conmensurabilidad de sus enseñanzas; y así sucesivamente hasta el último condicional. Para recomponer cada condicional se muestra que es necesario introducir una serie de elementos *mentales* para poder llegar a la definición y se desmantela la posibilidad de mantener la definición exclusivamente en términos de circunstancias y conductas observables objetivas.

Hemos visto hasta ahora las características de los distintos tipos de Conductismo que se desarrollaron durante la primera mitad del siglo XX. Todos ellos, aunque representan un avance en la investigación sobre la naturaleza de la mente, no pueden ser considerados teorías depuradas sobre la conciencia. Esto no implica que la tesis básica del Conductismo carezca por completo de valor y deba ser relegada al museo histórico de la filosofía de la mente. Por el contrario debemos reconocer en este esfuerzo una expresión de gran

importancia, que consiste en el hecho de que existe, sin lugar a dudas, una conexión conceptual entre las descripciones humanas sobre experiencias subjetivas con las descripciones teóricas de sus conductas. Resulta imposible comprender o explicar la conciencia evitando relacionar las disposiciones de conducta, las conductas mismas y las circunstancias observables con sus experiencias subjetivas asociadas. Ello posibilita concluir esta sección afirmando, al menos, que la conducta representa un conjunto de datos en tercera persona asociada a la explicación de la naturaleza de la conciencia. Que si bien no es todo lo que necesita explicación, al menos, si es introducida en un marco teórico acompañando la correlación con la experiencia subjetiva, ello puede conducir a una explicación más completa de la conciencia.

II.2. Teoría de la Identidad

En la época en que se desarrollaron las tesis conductistas no existía un conocimiento detallado sobre la manera en que determinados químicos actúan sobre el cerebro, alterando nuestros estados mentales (*e.g.* cómo las dopaminas atenúan el dolor, o el alcohol produce desinhibición o depresión) y en consecuencia se consideraba solamente que estas alteraciones son producto de una correlación entre la mente y el cerebro. Carnap postuló esta tesis como una versión anterior de la Teoría de la Identidad aceptada como una hipótesis empírica.

Por problema psicofísico no entendemos aquí la pregunta de si a todos los procesos psicofísicos corresponde un proceso simultáneo del sistema nervioso central (más precisamente si corresponde de tal manera, que a los procesos psíquicos semejantes pertenezcan procesos fisiológicos semejantes). Esto lo presuponemos aquí como una hipótesis empírica (Carnap, 1988, pág. 307).

El objetivo positivista que consistió en consolidar una ciencia universal de la naturaleza mediante un lenguaje de constitución, cuyas proposiciones básicas sirvieran de condiciones necesarias y suficientes, se mostró indiferente ante la posibilidad de justificar las bases de la psicología en una proposición empírica, como aquella que establece la identidad entre estados mentales y estados físicos. De acuerdo con la motivación principal, la psicología necesitaba estar cimentada, como las demás ciencias formales de la

naturaleza, en una verdad analítica a priori. No obstante en la medida en que las premisas en las que reposaban sus afirmaciones fueron revelando sus errores, la idea de que la psicología podría ser una ciencia sólida aun cuando únicamente podía apoyarse en proposiciones empíricas comenzó a incluirse en las investigaciones posteriores. La afirmación central de la Teoría de la Identidad es la más simple: los estados mentales *son* estados físicos del cerebro. Es decir, cada tipo de estado mental o proceso mental es una y la misma cosa que algún tipo de estado o proceso físico dentro del cerebro o del sistema nervioso central. El teórico de la identidad es paciente y confía en que la información que falta acerca del intrincado funcionamiento del cerebro se irá develando con los avances de la investigación.

En la actualidad existen razones suficientes para ofrecer un sinfín de enunciados de identidad en la ciencia; sabemos, por ejemplo, que el sonido es un tren de ondas de presión que se propagan por el aire, y que la propiedad de tener un tono alto, es idéntica a la propiedad de tener una frecuencia oscilatoria alta, que la luz son ondas electromagnéticas, ahora reconocemos que el calor o la frescura de un cuerpo es la energía del movimiento de las moléculas que lo componen; el calor es idéntico al alto valor medio de energía cinética molecular y el frío es idéntico al bajo valor medio de energía cinética molecular. Alguien con inclinaciones a defender la Teoría de la Identidad argumenta que un “estado mental”, es un proceso idéntico a un funcionamiento cerebral, exactamente de la misma forma que en los ejemplos anteriores y que la relación que guardan entre sí dichos enunciados de identidad puede ser descubierta mediante un proceso similar.

Si la motivación de la Teoría de la Identidad es explicar la naturaleza de la mente de la misma forma que se explican fenómenos físicos o químicos, es decir, mediante reducciones interteóricas, entonces es necesario encontrar datos sobre la mente que puedan ser reducidos a enunciados de identidad sobre procesos físicos objetivos. David Armstrong, en su artículo “La naturaleza de la mente” (Armstrong, 1980), remitió la pregunta sobre la naturaleza de la mente a la pregunta sobre la naturaleza del ser humano, señalando que el camino razonable para encontrar una respuesta sólida es explorar el conocimiento de la ciencia actual.

Los hombres tienen mentes, es decir, perciben, tienen sensaciones, emociones, creencias, pensamientos, propósitos y deseos. ¿Qué es tener una mente? ¿Qué es

percibir, sentir emoción, sostener una creencia o tener un propósito? Al igual que muchos otros filósofos modernos, pienso que la mejor clave que tenemos de la naturaleza de la mente nos la proporcionan los descubrimientos e hipótesis de la ciencia moderna que tiene que ver con la naturaleza del hombre (Armstrong, 1980, pág. 67).

De acuerdo con Armstrong la ciencia es capaz de ofrecer una explicación completa del hombre en términos puramente físicos, en términos elaborados por disciplinas científicas como la biología molecular, a través de desentrañar los mecanismos físicos y químicos que yacen en la base de la vida, o en términos de la neurofisiología que apunta a la posibilidad de una explicación electroquímica del funcionamiento del cerebro. En los términos que nos interesan para esta investigación hablamos de la transformación del problema mente-cuerpo en el problema mente-cerebro; desde la perspectiva de la Teoría de la Identidad implica en última instancia postular la ontología de la mente en términos físicos, es decir, que los procesos mentales deben distinguirse de (o relacionarse con) los procesos cerebrales en particular y no de los procesos corporales en general, puesto que implica la idea de que nuestras sensaciones, percepciones, recuerdos, inferencias, sentimientos y demás, tendrían lugar en nuestro cerebro como un proceso totalmente físico (sin la supuesta experiencia subjetiva asociada). Si la Teoría de la Identidad consiguiera explicar la mente en términos puramente físicos sería posible hablar de una explicación definitiva en términos de la objetividad deseada. De manera que de acuerdo con Armstrong es necesario tratar de elaborar una explicación de la naturaleza de la mente que sea compatible con la perspectiva de que el hombre no es sino un mecanismo físico-químico; y eso es precisamente de lo que se ocupa en su artículo.

Armstrong reconoce que los conductistas están ciertamente equivocados al identificar la mente (o las experiencias subjetivas) con la conducta, pero no completamente equivocados si atendemos a la conexión *lógica* que tienen dichos estados mentales con ella. Quizá la mente puede definirse no como conducta, sino como la *causa* interna de ésta. El reconocimiento de este éxito está ligado particularmente con la tesis conductista de Ryle que hemos esbozado en la sección anterior, de manera que no profundizaremos en ella de nuevo aquí. Al teórico de la identidad le parece bastante obvio que en los casos en los que pienso pero que de ello no resulta ninguna acción o conducta particular: “hay algo que está sucediendo en este momento, en el sentido más fuerte y literal de ‘suceder’, y este algo es

mi pensamiento” (Armstrong, 1980, pág. 69). De acuerdo con el teórico de la identidad la disposición de conducta de un individuo (o de potencia de reacción de un objeto, por llamarle de alguna manera) es en realidad un estado del objeto a reaccionar o ser propenso a comportarse de determinada forma bajo circunstancias determinadas, a diferencia de una sola propensión conductual. En la fragilidad del vidrio por ejemplo se trataría de un arreglo en la composición de sus moléculas; el hecho de que da lugar a las manifestaciones características de la fragilidad; es decir que la fragilidad del vidrio es idéntico al estado responsable de la propensión que tiene de manifestar su ruptura en circunstancias que causen que se rompa. Paralelamente en un estado mental se trataría del estado del cerebro que es el responsable del comportamiento característico; el hecho de que da lugar a las manifestaciones características de la percepción, por ejemplo, en circunstancias adecuadas. De acuerdo con Armstrong, si tratamos a las disposiciones propuestas por el Conductismo ryleano, como estados reales, incluso si son estados cuya naturaleza intrínseca necesita explicación (tales como las experiencias subjetivas), podemos decir que las disposiciones son causas reales que producen determinadas manifestaciones de conducta si se trata de estados mentales o de potenciales reacciones físicas si se trata de objetos físicos; en términos de estados mentales como causas de las sensaciones, percepciones y demás; o en el caso de lo que resulta menos misterioso como la fragilidad del vidrio, una cierta constitución molecular que constituye su fragilidad, responsable del hecho de que se rompa cuando es golpeado.

De esta manera los pensamientos, deseos, creencias y demás no se identifican con la conducta, su importancia en la explicación de la mente radica en el papel causal que guardan con la conducta, están detrás de ella; la conducta de un humano constituye la razón que tenemos de atribuirle ciertos procesos mentales, pero sin ser idénticos a estos. El pensamiento, por ejemplo no es el habla que se realiza bajo circunstancias adecuadas, más bien es algo dentro de la persona que, bajo circunstancias adecuadas, produce el habla. El hecho de que las disposiciones de conductas tengan en realidad un papel causal con respecto a las reacciones expresadas en estados mentales como percepciones, sensaciones, etc. es lo que reconoceríamos como los datos en tercera persona capaces de funcionar como agente de la objetividad deseada por la Teoría de la Identidad, ya que si las nociones de mente, estados mentales o experiencia subjetiva no son nada más que de causas dentro de la

persona, de ciertas respuestas conductuales, entonces se vuelve una preocupación científica mas no de análisis lógico la pregunta sobre la naturaleza intrínseca de esas causas. Suponiendo que esta caracterización de los estados mentales es correcta, de acuerdo con Armstrong, podemos identificar a los estados mentales con estados puramente físicos del sistema nervioso central. Concebir a la mente no como una conducta sino como aquello que la produce, representa un modo de superar los argumentos conductistas, incluso en el caso de las disposiciones, ya que de acuerdo con Armstrong en realidad se trata de estados que subyacen a ésta y que bajo circunstancias apropiadas, la producen. Un estado mental, concluye Armstrong, es un estado de la persona capaz de producir cierta gama de conductas. “La mente se concibe apropiadamente como un principio interno, pero un principio que se identifica en términos de la conducta externa que es capaz de producir bajo circunstancias determinadas” (Armstrong, 1980, pág. 71).

Con esta idea concuerda también David Lewis en su artículo “Un argumento en favor de la Teoría de la Identidad” (Lewis, 1984). El argumento de Lewis es sencillo: la característica definitoria de cualquier (tipo de) experiencia como tal es su papel causal, “su síndrome de las más típicas causas y efectos”. Si esos papeles causales pertenecen por necesidad analítica a las experiencias, entonces pertenecen de hecho a ciertos estados físicos. Por tanto puesto que esos estados físicos poseen las características definitivas de la experiencia, deben ser las experiencias. Considérese un ejemplo para abrir la imaginación: en los candados cilíndricos de combinación para cadenas de bicicleta, poner la combinación típicamente causa que se abran, es todo lo que necesitamos saber para atribuir que el candado está o no abierto. Podemos aprender que, como una cuestión de hecho, el candado contiene una fila de discos ranurados y que poner la combinación típicamente causa que las ranuras se alineen y al estar alineadas típicamente causa que el candado se abra. El alineamiento de las ranuras desempeña el papel causal que atribuimos al estar abierto por necesidad analítica, como la característica definitoria de estar abierto. Por lo tanto el alineamiento de las ranuras es idéntico con el estar abierto: ambos son uno y el mismo estado.

La Teoría de la Identidad, de acuerdo con Lewis, afirma que ciertos estados físicos son experiencias, procesos o actividades introspectivas, no que sean los supuestos objetos intencionales de los cuales las experiencias sean experiencias. Dice también que las atribuciones de experiencia tienen la misma referencia que ciertas atribuciones de estados

neuronales: ambas refieren por igual a los estados neuronales que son las experiencias. No dice que esas atribuciones tengan el mismo sentido; no lo tienen, las atribuciones de experiencia refieren a un estado especificando el papel causal que le pertenece accidentalmente en virtud de leyes causales mientras que las atribuciones de estados neuronales refieren a un estado describiéndolo en detalle. Lo esencial del argumento de Lewis es mostrar que los estados físicos que poseen las características definitivas de la experiencia, son la experiencia y que los correlatos que conocemos como la parte no física o la experiencia subjetiva asociada puede existir pero no tienen ninguna consecuencia causal, en realidad no debe ser tomada en serio porque no tiene ninguna consecuencia.

La primera de las dos premisas del artículo de Lewis para establecer la Teoría de la Identidad, es el principio de que la característica definatoria de cualquier experiencia es su papel causal. Esta premisa es una generalización de la teoría de J.J.C. Smart de que las manifestaciones de experiencia son de la forma “lo que me sucede es como lo que me sucede cuando...”; (e. g. lo que me sucede cuando siento dolor es lo que me sucede cuando mi cuerpo sufre algún daño), es decir, un tipo de tratamiento de una forma de expresarse común al de cualquier sujeto que cree que las experiencias son algo. Su origen es conductista en el sentido de que hereda el descubrimiento de que las conexiones causales entre una experiencia y sus manifestaciones típicas contienen necesidad analítica, pero permite que las experiencias sean algo real, a saber, los efectos de sus ocurrencias y las causas de sus manifestaciones. Permite también incluir otras experiencias entre las causas y efectos típicos mediante las que se define una experiencia. Recordemos aquí que para justificar la lista de condicionales necesaria para un análisis adecuado del lenguaje en el que se expresa una disposición, presuntamente constitutiva de todo estado mental, era necesario construir condicionales que incluyeran de cierto modo expresiones que denotaban experiencias y así el Conductismo no se mantuvo a salvo. El hecho de que exista esta posibilidad, hace justicia a la accesibilidad introspectiva. Por último, gracias a esta característica no estamos obligados a definir una experiencia por las causas y efectos de exacta y exclusivamente todas sus manifestaciones. Podemos, según Lewis, identificar la experiencia como aquel estado que típicamente es efecto de tales y cuales modos y típicamente causa de tales y cuales efectos, sin agregar nada acerca de éstos.

La segunda premisa del argumento de Lewis es la hipótesis de que hay una unidad de teorías científicas que proporciona un tratamiento verdadero de todos los fenómenos descriptivos en términos físicos. No tiene la intención de eliminar los supuestos fenómenos no físicos; no se trata de una tesis ontológica. Únicamente niega que necesitemos explicar los fenómenos físicos mediante fenómenos no físicos.

Los fenómenos físicos son explicables físicamente o son definitivamente inexplicables, en tanto dependen de la casualidad de una manera físicamente explicable o son metodológicamente aceptables como primitivos. Toda forma de fenómeno no físico puede coexistir con ellos al grado de compartir el mismo espacio tiempo, siempre y cuando los fenómenos no físicos sean enteramente ineficaces con respecto a los fenómenos físicos (Lewis, 1984, pág. 15).

La conclusión es que ninguno de esos fenómenos no físicos permisibles pueden ser experiencias; pueden estar ahí, como una clase separada de los que es objetivamente definible. Sin embargo todas las manifestaciones de conducta de las experiencias involucran o son ellas mismas fenómenos físicos y por tanto no pueden ser efectos de algo que no interactúa de ningún modo con el universo físico.

Esas manifestaciones de conducta están entre los efectos físicos definatorios de cualquier experiencia, de acuerdo con la primera premisa. Así, nada que sea ineficaz con respecto a los fenómenos físicos puede ser una experiencia. Por lo tanto, nada que sea un fenómeno no físico del tipo permisible bajo la segunda premisa puede ser una experiencia. De las dos premisas se sigue que las experiencias son uno y otro fenómeno físico (Lewis, 1984, pág. 16).

La argumentación de Lewis expone la duda sobre qué y cuáles fenómenos son aquellos que juegan el papel físico, y expresa incluso que la investigación está lejos de saberlo. En última instancia deposita la confianza, no en la explicación conceptual de los problemas filosóficos asociados a la mente, sino en las investigaciones y avances científicos. La naturaleza de lo que conocemos como conciencia, sus hechos físicos y sus supuestos elementos correlacionados con los eventos físicos que llamamos experiencia subjetiva, serán explicados con estos avances en algún punto en el futuro.

Existe también una preocupación implícita en el artículo de Lewis que es la necesidad de prevenir que el Dualismo encuentre una posición a salvo de la doctrina de que las experiencias son no-físicas y son físicamente ineficaces. Es verdad que tales fenómenos

nunca pueden ser refutados por ningún conjunto de teoría científica y evidencia. No obstante aun en esas condiciones, el problema se traduce en que no es posible que sean lo que llamamos experiencia. Únicamente puede ser los epifenómenos o correlatos no físicos de los estados físicos que son las experiencias (estados físicos que poseen las características definitivas de la experiencia). El Dualismo quedaría eliminado como una teoría capaz de ofrecer una explicación de la conciencia, de acuerdo con la tesis de Lewis, porque sucesos tales como los datos de la experiencia subjetiva pueden ser y las personas pueden sostenerlas en varios niveles, pero no tienen ninguna consecuencia; siguen siendo procesos privados. Si aceptamos que los enunciados acerca de los dolores, las punzadas, de cómo se ven las cosas, se oyen y se sienten, son enunciados que se refieren a sucesos y procesos internos, es decir en algún sentido privados, ¿nos estamos comprometiendo inevitablemente con alguna posición dualista, en la cual nuestras sensaciones e imágenes mentales forman parte de una categoría diferente de la que forman parte los procesos físicos, con los que se supone están correlacionados? La respuesta es no de acuerdo con Ullin Place en su artículo “¿Es la conciencia un proceso cerebral?” (Place, 1956).

No se sostiene que los enunciados acerca de las sensaciones y de las imágenes mentales sean reducibles a enunciados acerca de procesos cerebrales, ya que decir que los enunciados acerca de la conciencia son afirmaciones acerca de procesos cerebrales es evidentemente falso, por lo siguiente:

a) Uno puede describir sus sensaciones e imágenes mentales sin saber nada acerca de sus procesos cerebrales, incluso sin saber si éstos existen. (e.g uno puede describir la experiencia sensorial y emocional desagradable de dolor asociada con una lesión, sin saber nada al respecto de, por ejemplo, la teoría de la especificidad, que se basa en la existencia de terminaciones nerviosas libres que vinculan la información receptiva con el sistema nervioso central).

b) Las afirmaciones acerca de la conciencia de uno mismo y las afirmaciones acerca de los procesos cerebrales de uno mismo, se comprueban de manera totalmente diferente.

c) El hecho de que no hay nada auto-contradictorio en la afirmación “X tiene un dolor pero no está pasando nada en su cerebro”.

Se supone entonces que la afirmación “la conciencia es un proceso cerebral” no es ni auto-contradictoria ni autoevidente, sino que es una hipótesis científica razonable, de la misma manera en que lo es la afirmación “la energía molecular promedio es proporcional a la temperatura”. Existen buenas razones para que el teórico de la identidad crea que la neurociencia llegará a cumplir las condiciones necesarias para anunciar que ha encontrado la identidad de los estados mentales y los estados neuronales. Pero también se encuentra con dificultades teóricas cuando intenta explicar la conciencia cabalmente.

II.2.1. Argumentos a favor de la Teoría de la Identidad

Orígenes físicos

Los individuos humanos comenzaron por ser una organización monocelular de moléculas programadas genéticamente, y a partir de allí se produce un desarrollo mediante la adición de más moléculas cuya estructura e integración está controlada por la información codificada en las moléculas de ADN. La consecuencia de este proceso sería un sistema puramente físico cuya conducta es el resultado de su funcionamiento interno y de sus interacciones con el resto del mundo físico. Y precisamente aquello de lo que se ocupan las neurociencias son esas operaciones internas que controlan la conducta.

La evolución de procesos físicos

La teoría de la evolución constituye la única explicación seria que tenemos para dar cuenta de la capacidad del cerebro y del sistema nervioso central para controlar la conducta. En pocas palabras, estos sistemas fueron seleccionados por las múltiples ventajas (en última instancia, la ventaja de la reproducción) que otorga a las criaturas cuya conducta se controla de esta manera. Una vez más en este caso parecería que las causas básicas de nuestra conducta se remiten a la actividad nerviosa.

Dependencia del sistema nervioso.

Este argumento, descrito anteriormente, surge a partir del éxito de las neurociencias en la tarea de describir el sistema nervioso de muchos seres y explicar sus aptitudes y deficiencias conductuales en términos de las estructuras descubiertas.

II.2.2. Argumentos en contra de la Teoría de la Identidad

Argumento de la introspección

La Teoría de la Identidad comparte una cierta debilidad con el Conductismo: es una explicación satisfactoria en términos de la objetividad deseada, de los datos en tercera persona; desde el punto de vista de lo que es públicamente observable. Sin embargo no funciona como una explicación desde la perspectiva de la primera persona; en los casos en los que quisiéramos explicar nuestra propia corriente de conciencia, parece que nos damos cuenta de muchas más cosas que sólo de nuestra conducta. En la introspección se nos revela un ámbito de pensamientos, sensaciones y emociones, no de impulsos electroquímicos en una red nerviosa o una simple conducta. No podemos decir que estar conscientes y tener experiencias es simplemente que algo suceda dentro de nosotros capaz de causar ciertos tipos de conducta. Los estados y propiedades mentales que pone de manifiesto la introspección parecen algo radicalmente distinto de los estados y propiedades neurofisiológicos o conductuales, y que evidencian interacciones causales por sí solos. Sin más insistiremos en que sin una explicación de esta parte de la naturaleza de la mente, la explicación sobre la conciencia está seriamente incompleta.

Errores categoriales

La identificación de estados mentales con estados cerebrales nos llevaría a afirmaciones literalmente ininteligibles, a lo que los filósofos han denominado “errores categoriales”. Una de las leyes más importantes sobre la identidad numérica es la Ley de

Leibniz, esta ley postula que dos ítems son numéricamente idénticos sólo en caso de que cualquier propiedad que postule uno de ellos la posea también el otro. En notación lógica:

$$(\forall x)(\forall y) [(x=y) \equiv F (Fx \equiv Fy)]$$

Esta ley apunta hacia un modo posible de refutar la Teoría de la Identidad: encontrar una propiedad que poseyeran los estados cerebrales pero no los estados mentales (o viceversa) desacreditaría la teoría. Al respecto de esta dificultad se ha apuntado que los estados y procesos cerebrales deben tener alguna localización espacial específica; en el cerebro o en alguna parte. Si los estados mentales son idénticos a los estados cerebrales entonces deben tener exactamente la misma localización espacial; lo que parece que es un error, ya que el Dualismo postula que los estados mentales no poseen ni forma, ni masa, ni posición espacial y hasta el momento no parece haber forma de refutar esa postura.

II.3. Funcionalismo

El Funcionalismo es una de las propuestas más importantes que se han ofrecido como explicación al problema mente-cuerpo. Las soluciones anteriores han basado sus esfuerzos en responder preguntas como ¿Cuál es la naturaleza última de lo mental? ¿Qué es lo que hace un estado mental y su experiencia subjetiva, un estado mental? Para responder estas preguntas el Dualismo propuso que la naturaleza última de lo mental se halla en una propiedad especial no física. El Conductismo señaló que los estados mentales se identifican con disposiciones conductuales y la Teoría de la Identidad, en su versión más influyente, apunta que dicha identificación se da entre estados mentales y estados físicos cerebrales.

El Funcionalismo, en cambio, es la tesis de que los estados mentales son estados funcionales, es decir, que si una propiedad funcional F pertenece a un objeto es debido exclusivamente a la aptitud que tiene el objeto de cumplir cierto papel causal en un determinado contexto. Así que el rasgo esencial de todo tipo de estado mental es el conjunto de relaciones causales que mantiene con los efectos ambientales sobre el cuerpo, otros estados mentales y la conducta. Así que las preguntas dejan del lado la necesidad de

responder cuestiones sobre la naturaleza de lo mental o de sus correlatos conductuales o sus reducciones, y comienza a ser preguntas acerca de sus causas y efectos.

Por ejemplo, el rasgo característico de un dolor es que es el resultado de alguna lesión corporal que provoca angustia, incomodidad y algún tipo de razonamiento práctico destinado a aliviarlo. También genera que el individuo procure cuidados a la zona afectada. Así que cualquier cosa que cumpla exactamente con esta función, es para el funcionalista, un dolor. El caso de otros estados mentales como las sensaciones de cosquilleo, la ira, el amor, la sensación de oscuridad, creer que hoy es lunes y desear que fuera viernes, es similar; su definición se da en virtud de las funciones causales específicas que son intermediarios entre los *inputs* de estímulos sensoriales y los *outputs* en forma de conductas.

Quizá esta caracterización nos recuerde al Conductismo que identificaba estados mentales con disposiciones a actuar de ciertos modos en ciertas circunstancias de *input*. Pero de acuerdo con el Funcionalismo en un caso en el que S desea P como un objetivo, no puede identificarse P con la disposición A en condiciones en las que A ocasione que P, dado que S puede perfectamente no saber que A ocasiona P y por tanto no estar dispuesto a realizar A. Recordemos que para el conductista es necesario y suficiente para autentificar un estado mental como desear P, que S realice un conjunto de condiciones que señalen su inclinación por P. En cambio desde la visión funcionalista, un sistema, consciente o no, puede manifestar relaciones *input-output* aunque no desee P, porque el hecho de que tal sistema desee que P depende de que tenga estados mentales (y en última instancia experiencias subjetivas) que tienen relaciones causales con otros estados mentales. En este punto resulta muy arriesgado afirmar esto, pero concédanse la verdad de estas ideas, con la garantía de que quedará clarificado en el transcurso de esta sección.

Cabe destacar un elemento importante, para el Funcionalismo no es importante el material del que esté constituido fisiológicamente un *sistema*. Dos sistemas, cuya constitución química y física fuera distinta, podrían mantener las mismas funciones causales específicas. Ambas podrían tener un “estado interno” que cumpla con todas las condiciones para ser un estado de dolor como el de la caracterización anterior. Esto se conoce como el principio de realizabilidad múltiple.

Examinemos esta afirmación mediante el camino que muchos autores han seguido para dar cabida a esta particularidad, además de ser el más pertinente para nuestro análisis. Nos referimos a lo que muy naturalmente ahora se conoce como Máquina de Turing que representa, a nuestro modo de ver, la versión más efectiva del Funcionalismo, y por tanto la que vale más discutir en este trabajo.

Una Máquina de Turing no es una máquina en el sentido ordinario de la palabra, es decir, no es una máquina física, sino una especificación abstracta de una posible máquina que formaliza la idea de lo que significa computar algo¹⁴. Una máquina de Turing simple tiene las siguientes características: una cinta dividida en cuadrados y un dispositivo que puede escribir símbolos en la cinta y luego leer tales símbolos. El dispositivo se encuentra alternativamente en ciertos “estados internos” y puede mover la cinta a la derecha o a la izquierda un cuadro a la vez. Consideremos que hay sólo dos clases de símbolos que pueden ser escritos en la cinta expresados en los numerales “0” y “1”, cada símbolo ocupa un cuadrado de la cinta. De este modo el dispositivo puede hacer cuatro cosas:

1. Puede mover la cinta un cuadrado a la vez, de izquierda a derecha o de derecha a izquierda.
2. Puede leer un símbolo en la cinta.
3. Puede escribir un símbolo en la cinta, ya sea escribiendo en un cuadro en blanco o encima de otro símbolo.
4. Puede cambiar su “estado interno”

Las operaciones que la máquina pueda resolver se representan por medio de su tabla de máquina, una serie de instrucciones de la forma “si la máquina está en el estado X y leyendo el símbolo S, entonces realizará determinada operación (por ejemplo, escribir o borrar un símbolo, moviendo la cinta) y pasará al estado Y (o seguirá en el mismo estado) y moverá la cinta a la derecha/izquierda”. Se puede pensar en estas instrucciones como el “programa” de la máquina. Al especificar una posición particular en la tabla de máquina, necesitamos saber dos cosas: la entrada del momento a la máquina y su estado del

¹⁴ Google creó un divertido doodle que ilustra de manera lúdica el funcionamiento de una Máquina de Turing. Se puede acceder a él aquí: <http://www.google.com/doodles/alan-turings-100th-birthday>.

momento. Lo que la máquina hace es enteramente fijado por estas dos cosas. Imaginemos una Máquina particular de Turing que compute únicamente las siguientes instrucciones, para facilitar la comprensión:

1. Si la máquina está en el estado A, y lee un 0, entonces permanece en el estado A, escribe un 0 y se mueve un cuadrado a la derecha.
2. Si la máquina está en el estado A, y lee un 1 entonces cambia al estado B, escribe un 1 y se mueve un cuadrado a la derecha.
3. Si la máquina está en el estado B, y lee un 0, entonces cambia al estado A, escribe un 1 y se detiene.
4. Si la máquina está en el estado B, y lee un 1, entonces permanece en el estado B, escribe un 1 y se mueve un cuadrado a la derecha.

Una tabla de máquina para esta particular Máquina de Turing sería como sigue:

		ENTRADA	
		1	0
ESTADO DE LA MÁQUINA	A	Cambiar a B; Escribir un 1; Mover la cinta a la derecha	Estar en A; Escribir un 0; Mover la cinta a la derecha
	B	Estar en B; Escribir un 1; Mover la cinta a la derecha	Cambiar a A; Escribir un 1; ALTO

Una Máquina de Turing descrita de esta manera es cierto tipo de dispositivo con un funcionamiento de *input/output*. De acuerdo con Ned Block (Block, 1996) esta ilustración sirve para puntualizar algunas afirmaciones del Funcionalismo.

El Funcionalista afirma que la naturaleza de un estado mental es como la naturaleza de un estado de la máquina anterior. A es estar en un estado tal que, al recibir una entrada (*input*) de “1” ocasiona tal y cual cosa; por ejemplo, una definición funcionalista de la teoría psicológica según la cual un dolor está causado por un daño en el cuerpo que causa preocupación, la emisión de un “¡Auch!” y de que la preocupación causa a su vez el fruncimiento de las cejas sería como sigue:

‘x’ tiene (siente) dolor si y sólo si hay dos estados, el primero de los cuales está causado por daño en el cuerpo y a su vez causa la emisión de un “¡Auch!” y el segundo estado, y el segundo estado causa fruncir el entrecejo, y x está en el primer estado.

El método de la Máquina de Turing funciona para cualquier estado mental, de acuerdo con Block, y cumple además una de las condiciones deseadas por el Conductismo: caracterizar lo mental en términos completamente no mentales; en este caso sólo de funciones lógicas o matemáticas. A es un estado de segundo orden que implica tener una clase de propiedades, por ejemplo mecánicas, hidráulicas o electrónicas. De manera que un estado mental caracterizado en términos funcionales implica que puede ser realizado de diferentes maneras; en el caso de A sería un estado mecánico, si la máquina estuviera hecha de metal y plástico. Además un estado físico puede realizar diferentes estados funcionales en diferentes sistemas (conscientes o no). De modo que si una criatura sin cerebro como la Máquina de Turing puede realizar tal economía funcional, dicho proceso no puede ser un estado cerebral, sino que estaría caracterizado en términos de su estructura objetiva de datos en tercera persona.

Existe al menos una relación evidente entre lo que hemos señalado acerca de la mente y lo que hemos averiguado acerca de la Máquina de Turing. La noción de representación parece figurar en ambas áreas. Como hemos analizado en el capítulo anterior, uno de los rasgos esenciales de ciertos estados mentales es que representan. Y hemos visto que una Máquina de Turing puede, en cierto sentido, procesar representaciones; permite describir computacionalmente los estados mentales, realizándose en una gran variedad de materiales.

Lo que ilustra esto es la posibilidad de que la naturaleza (y el hombre) tenga más de un camino para diseñar un ser que piense, sienta, perciba y actúe. Es posible que existan muchos tipos diferentes de sistemas físicos que puedan realizar la economía funcional característica de la inteligencia consciente. Paul Churchland nos invita a imaginar un

ejemplo, agregando que: “Lo que cuenta en el terreno de lo mental no es la materia de la que está hecho un ser, sino la estructura de las actividades internas que sostiene esa materia” (Churchland, 1984/1999, pág. 66).

Imaginemos a un ser de otro planeta, dice el funcionalista, un ser con una constitución fisiológica distinta, que se basa en un elemento químico como el silicio, por ejemplo, en lugar de estar basada en el carbón como la nuestra. La estructura química, e inclusive la estructura física del cerebro de este ser extraño tendría que ser sistemáticamente diferente de la nuestra. Pero, aun así, el cerebro de ese ser bien podría mantener una economía funcional de estados mentales cuyas relaciones mutuas se correspondieran perfectamente con las relaciones mutuas que definen los nuestros. El extraterrestre puede tener un estado interno que cumpla todas las condiciones para ser un estado de dolor, como hemos explicado antes. Ese estado, considerado desde un punto de vista puramente físico, tendría una estructura muy diferente de la del dolor humano, sin embargo podría ser idéntico a un estado de dolor humano desde un punto de vista puramente funcional. Y lo mismo podría decirse para todos sus estados funcionales (Churchland, 1984/1999, págs. 65-66).

Entonces es lógicamente posible imaginar la organización funcional de un sistema consciente en diferentes variedades de sistemas físicos, además del humano: extraterrestres, robots, zombis, ordenadores etc. De acuerdo con John Searle (Searle J. , 1985), uno de los puntos más populares en filosofía, psicología e inteligencia artificial es aquel que compara el funcionamiento del cerebro con el funcionamiento de una computadora digital¹⁵. En la versión más extrema de esta analogía, el cerebro *es* una computadora digital y la mente su sistema operativo. En esta afirmación no hay nada esencialmente biológico que sirva de soporte para la mente. El cerebro humano es sólo uno de un número indefinidamente extenso de posibles *hardwares* capaces de alojar una organización funcional de la mente o su *software*. Cualquier sistema cargado con la aplicación correcta cuya función sea el procesamiento de *inputs* interactivos con el ambiente y otros *softwares* en forma de experiencias perceptuales, corporales y emocionales, imágenes mentales y pensamientos

¹⁵ Como un correlato histórico al respecto de esta afirmación el mismo Searle apuntó en su libro “Mentes, cerebros y ciencia” que: “Porque no comprendemos el cerebro muy bien, estamos constantemente tentados a usar la más reciente tecnología como modelo para tratar de comprenderlo. En mi infancia siempre se aseguraba que el cerebro era un tablero telefónico... Sherrington, el gran neurocientífico británico, pensaba que el cerebro funcionaba como un sistema telegráfico. Freud a menudo comparó el cerebro con sistemas hidráulicos y electromagnéticos. Leibniz lo comparó con un molino, y se me ha dicho que algunos de los antiguos griegos pensaban que el cerebro funcionaba como una catapulta. Hoy por hoy, evidentemente, la metáfora es la computadora digital” (Searle J. , 1985, pág. 44).

ocurrentes, y con los *outputs* en forma de conductas, presumiblemente tendría una mente en el mismo sentido que cualquier ser humano. Searle cita a Allan Newell afirmando que:

Hemos descubierto (y obsérvese que Newell dice ‘descubierto’ y no ‘hemos avanzado la hipótesis’ o ‘hemos descubierto la posibilidad’, sino hemos *descubierto*) que la inteligencia es solamente un asunto de manipulación de símbolos físicos; no tiene ninguna conexión esencial con ningún género específico de *wetware* o hardware biológico o físico (Searle J. , 1985, pág. 35).

Y a John McCarthy agregando que:

Toda máquina capaz de resolver problemas puede decirse que tiene creencias [...] puede decirse que máquinas tan simples como los termostatos tienen creencias [...] Mi termostato tiene tres creencias: ‘hace demasiado calor aquí’, ‘hace demasiado frío aquí’ y ‘aquí hace la temperatura correcta’ (Searle J. , 1985, págs. 35-36).¹⁶

Los ejemplos anteriores pueden llevar a la imaginación a lugares insospechados pero existe un problema con ese tratamiento.

La noción de procesamiento de información está siendo tratada con mucha ligereza en este caso. Ciertamente un estado mental como el pensamiento implica procesar información; cuando pensamos, obtenemos información del medio ambiente y la usamos para conducirnos en el mundo. Sin embargo no es un camino natural pasar de esto y de que las máquinas procesen información, a la conclusión de que dicho procesamiento de información que realicen las máquinas sea un género de pensamiento o el pensamiento mismo. La cuestión es entonces que no se puede apuntar que la proposición “procesamiento de información” puede aplicarse tanto a las computadoras como al pensamiento humano de la misma forma: esto es una falacia de confusión.

En una computadora digital, o un termostato ‘creativo’, es necesario especificar procesos a los artefactos en términos de lenguajes formales; en códigos binarios, por ejemplo. Introducidos en una computadora los símbolos en forma de numerales ‘0’ y ‘1’ no

¹⁶ Cabe recordar que en algún sentido el cerebro humano es una máquina, un sistema físico que es capaz de realizar cierto género de operaciones, una máquina de carne; inductivamente podemos afirmar de esta única premisa que hay máquinas que pueden pensar.

tienen significado en sí mismos, no son número como tal, ni tienen contenido semántico, son sólo datos que no se refieren a nada. Su característica principal es que se describen en términos únicamente de su estructura sintáctica y en ello radica su poder; puede ser introducido en diferentes sistemas y usarse para implementar diferentes respuestas. Pero si se utilizan como corolarios de los procesos mentales, el resultado falla, y la razón es simple: tener una mente es algo más que tener procesos formales sintácticos. Las experiencias subjetivas internas tienen por definición, cierto tipo de contenido; los datos de la mente no son sólo los rasgos formales objetivos en tercera persona, sino que poseen además datos con contenido semántico.¹⁷

Searle propone el siguiente argumento que arroja varias conclusiones al respecto:

1. Los cerebros causan las mentes
2. La sintaxis no es suficiente para la semántica
3. Los programas de computadora están definidos enteramente por su estructura formal o sintáctica.
4. Las mentes tienen contenidos mentales; específicamente, tienen contenidos semánticos

Conclusión 1: ningún programa de computadora es suficiente por sí mismo para dar un sistema, una mente. Los programas, dicho brevemente, no son mentes, y no son suficientes por sí mismos para tener mentes.

Conclusión 2. El modo en que las funciones del cerebro causan las mentes no puede ser solamente en virtud de pasar un programa de computadora.

Conclusión 3: Cualquier otra cosa que cause las mentes tendría que tener poderes causales equivalentes al menos a los del cerebro.

Conclusión 4. Para cualquier artefacto que pudiéramos construir que tuviese estados mentales equivalentes a los estados mentales humano, el desarrollo de un programa de

¹⁷ Véase el conocido ejemplo de la habitación china en (Searle J. , 1985, págs. 33-48)

computadora no sería suficiente por sí mismo. Más bien, el artefacto tendría que tener poderes equivalentes a los del cerebro humano.

De acuerdo con Searle el problema asociado con la experiencia consciente “El problema de la conciencia en singular” es el problema de explicar exactamente cómo los procesos neurobiológicos en el cerebro causan nuestros estados subjetivos; cómo exactamente esos estados son realizados en/por los estados cerebrales. Apuntando además que “[...] la razón de que la conciencia parezca un misterio es que no tenemos una idea clara del modo en que algo en el cerebro pueda causar estados conscientes”. (Searle J. R., 2000, pág. 13). De manera que el argumento anterior le permite a Searle afirmar que:

El resultado de esta discusión es recordarnos algo que ya sabíamos desde el principio, a saber: que los estados mentales son fenómenos biológicos. La conciencia, la intencionalidad y la causación mental son todas ellas parte de la historia de nuestra vida biológica, junto con el crecimiento, la reproducción, la secreción de la bilis y la digestión (Searle J. , 1985, pág. 37).

A nuestro modo de ver, la discusión hasta este punto expone un evidente género de Dualismo residual, que consideramos por dos razones: a) porque el Funcionalismo descansa sobre el rechazo de la idea de que la mente es sólo un fenómeno biológico natural del mundo igual que cualquier otro, b) porque no se ha ofrecido una explicación sobre la naturaleza misma de los contenidos mentales semánticos; en cierto modo la razón de que no se haya hecho es porque presumiblemente había resultado irrelevante para las posiciones filosóficas anteriores. Pues bien, ocupémonos de la primera cuestión. Existe una cercanía sospechosa entre el Dualismo cartesiano y el Funcionalismo de Máquina de Turing por una razón no tan evidente. Mientras el primero afirma que es concebible/posible separar mente y cuerpo, el segundo sostiene que no es estrictamente necesario que los estados mentales se realicen en las propiedades físicas de engranajes y máquinas reales o en las propiedades físicas de un cerebro.

El poder de una Máquina de Turing radica, insistimos, en que puede en principio implementar cualquier procedimiento efectivo calculable y esto presumiblemente es crucial para la futura construcción de máquinas programadas capaces de exhibir la misma inteligencia de un humano. Algunos defensores de la Inteligencia Artificial fuerte aseveran que la mente puede modelarse computacionalmente, ya que una máquina que opere como

un cerebro *es* un cerebro artificial. La mente humana sería de este modo un computador digital apropiadamente programado y en consecuencia hablaríamos de estructuras simbólicas dentro del cerebro que se correlacionan y explican las actividades del pensamiento y de la inteligencia humana. Pero el tema de la Inteligencia Artificial es demasiado amplio para introducirlo en esta revisión del Funcionalismo, así que la dejaremos de lado por ahora. No obstante nos guiaremos de ella para introducir la problemática asociada del Dualismo residual que encontramos en el Funcionalismo. Searle afirma en diversas ocasiones que:

La IA es una extraña mezcla de Conductismo y Dualismo. Es conductista en su aceptación del funcionamiento de una Máquina de Turing, pero en un nivel más profundamente filosófico es dualista, porque rechaza la idea de que la conciencia y la intencionalidad son fenómenos biológicos ordinarios como la digestión [...] debemos pensar en la mente como “un tipo de cosa abstracta cuya identidad es independiente de la encarnación física. (Searle J. , 2002, pág. 57)

En conclusión el Funcionalismo de Máquina de Turing no *implica* una forma de Dualismo por sí sola, pero dado que es el paradigma del principio de realizabilidad múltiple sí resulta compatible con una de sus formas.

La segunda problemática interesante es la falta de explicación de los contenidos mentales semánticos. ¿A qué se refiere el Funcionalismo cuando habla de contenidos mentales semánticos? En una posible respuesta Sidney Shoemaker señala que: “En una de sus interpretaciones, el Funcionalismo en la filosofía de la mente es la doctrina de que los términos mentales o psicológicos son en principio eliminables de cierta manera” (Shoemaker, 2003) Nótese que Shoemaker está hablando de los términos, lo que a primera vista señala que se trata de las problemáticas asociadas al conflicto semántico de la mente. Los funcionalistas han tratado a los estados mentales como identificaciones de la totalidad de los estados físicos de un sistema, se definen a sí mismos mediante una mención explícita de los *inputs* y *outputs* en lenguaje físico. El Funcionalismo reconoce a los estados mentales únicamente mediante la especificación física, o al menos no mental, de *inputs* y *outputs* en un sistema. Mediante un acercamiento básico a los problemas asociados al Conductismo como una buena explicación de la conciencia en su totalidad, reconoceríamos que el Funcionalismo no puede evitar el tipo de problemas por los cuales el Conductismo fracasó, a saber, la problemática asociada a la experiencia subjetiva.

Al respecto del problema Hilary Putnam intentó mostrar que con el modelo de la Máquina de Turing, la pretendida naturaleza única de la experiencia subjetiva no podía entrar en consideración y que desaparecería con el nuevo panorama del Funcionalismo. De tal modo que si consideramos la Máquina de Turing (en general: dispositivos mecánicos que reciben *inputs*, que aplican reglas algorítmicas, ‘computan’, y emiten *outputs* verbales y conductuales), no hay enigma alguno en lo que pudiera relacionar un dispositivo mecánico con su programa; entre los estados estructurales y los estados lógicos, sintácticos y algorítmicos de un sistema. Por lo tanto tampoco tendría que haber misterio, a la luz de una analogía completa, en lo que se refiere a la relación entre el cerebro y los estados mentales de una persona. Quizá deberíamos comenzar la discusión sobre las almas de las computadoras o en definitiva abandonar por salud, la discusión sobre el alma de las personas; pero ese es otro tema.

De acuerdo con Putnam (Putnam, 1981) la teoría de que un dolor es una disposición de conducta posee una ventaja aparente frente al Funcionalismo: parece concordar con la manera como verificamos que los organismos tienen experiencias subjetivas. No hay nada en la práctica corriente acerca de que cuando un animal siente, percibe o piensa se encuentra en un estado cerebral, y además, tenemos poco conocimiento de su organización funcional. Pero esta ventaja es sólo aparente, ya que decir que una experiencia subjetiva no es ni un estado cerebral ni un estado funcional, es semejante a argumentar que el calor no es la energía cinética molecular media de la reducción interteórica, que propone la Teoría de la Identidad, partiendo del hecho de que las personas no descubren la energía cinética molecular cuando investigan sobre lo caliente o frío de algo; no es necesario. Lo que sí es necesario es que los signos que toma como indicaciones de calor sean de hecho explicados por la energía cinética molecular media. Del mismo modo la hipótesis funcionalista requiere que los signos que se toman como indicaciones conductuales de la experiencia subjetiva se expliquen a partir del hecho concreto de estar en un estado funcional, aunque los hablantes no sepan nada al respecto.

Consideremos, por su parte, lo que el teórico de la identidad tiene que especificar para que sus afirmaciones sean efectivas. Tiene que especificar un estado fisicoquímico tal que cualquier organismo (no solamente un mamífero) puede sentir, percibir, o pensar si y sólo si posee un cerebro con una estructura fisicoquímica adecuada y que su cerebro esté en ese

estado físico-químico. Esto significa que dicho estado debe necesariamente ser un estado posible en el cerebro de un mamífero, un reptil, un molusco, etc. Y al mismo tiempo que no sea posible en el cerebro de cualquier otra criatura incapacitada naturalmente para sentir, percibir o pensar. El teórico de la identidad extiende la misma naturaleza, para cualquier estado psicológico. En contraste el Funcionalismo identifica el dolor, el hambre, la ira, el calor y demás experiencias subjetivas sobre la base de los efectos causales de los estímulos ambientales, la conducta y la respuesta funcional que se deriva de ellos. Así cualquier animal que consideremos capaz de mostrar dicha diferencia, por lo menos parecerá tener una cierta clase de organización funcional.

II.3.1. Argumentos en contra del Funcionalismo

La objeción que más comúnmente se la ha formulado al Funcionalismo tiene que ver con de la experiencia subjetiva, precisamente con los qualia. Parece que el Funcionalismo puede superar una de las dificultades que no superaba el Conductismo filosófico, no obstante, cae presa de la otra. Al considerar como rasgo definitorio de todo estado mental a sus propiedades relacionales, el Funcionalismo ignora su naturaleza interna, o cualitativa, que es el rasgo definitorio de muchos tipos de estados mentales (i.e. el dolor, la sensación de color, de temperatura, etc.) según lo expresa esta objeción. Por tanto el Funcionalismo es falso.

Posibilidad del espectro invertido

Es lógicamente posible que la gama de sensaciones de color que un sujeto experimenta, cuando observa objetos comunes, esté invertida en relación con la sensación de color que experimenta otro sujeto. Dos sujetos que conducen un auto, al mirar en el semáforo la luz que indica alto, desean detenerse y pisan el freno. Al ver la luz es lógicamente posible que uno de ellos experimente lo que “normalmente” es la sensación de rojo, mientras que el otro, experimente la sensación “normal” de amarillo; sin importar que ambos la llamen roja. Aun si esto ocurre, ambos sujetos siguen siendo mutuamente isomórficos desde el punto de vista funcional. La sensación visual de ambos ante el color

del semáforo que indica alto, es funcionalmente idéntica. Constituye el mismo tipo de estado y no tiene sentido preguntarse si su sensación es, o no, la sensación “real” de rojo. Este tipo de posibilidades lógicas ocurren de hecho, concluye el argumento, por tanto si el Funcionalismo las niega, entonces es falso.

Qualia ausentes

Es lógicamente posible imaginar la organización funcional de un sistema consciente en diferentes variedades de sistemas físicos, además del humano: ordenadores, extraterrestres, robots, zombis, etc. Todos ellos poseen en cierto grado la objeción de los qualia ausentes, no obstante, ninguno lo plantea mejor que el ejemplo de la nación china: imaginemos a la totalidad de la población de china; a los 1 331 460 000 chinos organizado en una compleja red de interacciones mutuas que les permitiera constituir un cerebro gigantesco, inserto en un cuerpo robótico individual, con su respectivo intercambio de *inputs* sensoriales y *outputs* conductuales. La unidad cerebral compuesta por la población china mas la estructura física robótica, presumiblemente ejemplificaría la organización funcional pertinente, y por tanto sería el sujeto de estados mentales, de la forma que propone el Funcionalismo. Pero los estados complejos que ahí cumplen las funciones de dolor, placer y sensaciones de color, seguramente no tendrían qualia internos como los que poseemos los seres humanos, por lo tanto no pueden ser auténticos estados mentales. Así que el Funcionalismo está incompleto.

El esfuerzo materialista que estudió la comprensión de procesamientos objetivos en términos de datos en tercera persona funcionales, presenta un modo de explicar la conciencia humana. Sin embargo queda incompleto, dado que no propone un tratamiento concluyente sobre la naturaleza de los datos en primera persona y la experiencia subjetiva, supuestamente asociada a dichos mecanismos objetivos. Insistiremos en lo sucesivo, en que la explicación de la conciencia en su totalidad requiere de la integración de los datos de la experiencia subjetiva en un mismo marco teórico de investigación y esperar que dicha integración destile de su objeto de estudio las ideas confusas del mentalismo. En lo sucesivo veremos cómo es posible este panorama.

II.4. Materialismo Eliminativo y Naturalismo Mecanicista

El Materialismo Eliminativo, o simplemente Eliminativismo,¹⁸ es la tesis de que nuestra concepción de sentido común acerca de los fenómenos psicológicos conocidos popularmente constituye una teoría radicalmente falsa, una teoría tan defectuosa que tanto sus principios como su ontología serán desplazados sin más por una neurociencia completa, en lugar de ser reducidos. En principio el materialismo eliminativo se vio motivado por la puesta en duda de la Teoría de la Identidad, ya que parecía poco probable que sus tesis trajeran consigo las correspondencias reductivas entre los fenómenos físicos y sus correlatos neuronales; o entre los conceptos de la psicología popular y los conceptos de la neurociencia teórica. La razón de ello es lo que hemos analizado anteriormente, la posibilidad de la diversidad de sistemas físicos capaces de sostener la organización funcional requerida. El Eliminativismo posee paralelos históricos que apoyan sus afirmaciones. Así como la Teoría de la Identidad ha logrado mostrar ejemplos de una reducción interteórica exitosa, el Eliminativismo es capaz a su vez de mostrar que hay buenos ejemplos en los que se produjo la eliminación de la ontología de una teoría anterior y la consecuente sustitución por una teoría con mayor poder explicativo.

En el siglo XVII Ernst Stahl desarrolló la idea de que cuando una pieza de madera se quema o una barra de metal se oxida, se liberaba una sustancia inmaterial llamada “flogisto”; a gran velocidad en el primer caso y con mucha lentitud en el segundo. Una vez que se liberaba la sustancia por completo, dejaba tras de sí una pila de ceniza o herrumbre carente de flogisto. Pero con la investigación posterior, Antoine Lavoisier demostró que esta interpretación de la combustión era falsa ya que se trataba de una reacción química particular. En ambos casos se logró comprender que no se trataba de la pérdida de algo sino del agregado de una sustancia presente en la atmósfera: el oxígeno. Así que la idea del flogisto reveló ser, no sólo una interpretación incompleta, sino una descripción totalmente equivocada de un fenómeno de la naturaleza, así que quedó eliminada como explicación

¹⁸ Consideramos que en español el término correcto es “Eliminacionismo” ya que “Eliminativismo” proviene de la locución del inglés “eliminative”, no obstante no usaremos, en este trabajo, el término correcto en español, dado que en la literatura filosófica que aborda el tema es más común encontrar el segundo término.

científica concluyente. Esta fórmula es la intuición sobre la que se centra la tesis del Eliminativismo en filosofía de la mente.

¿Qué tendría que ser eliminado o sustituido en última instancia en lo relativo a la mente y el cerebro? ¿Qué ha ocurrido con la ontología de los datos en primera persona en su relación con los datos en tercera persona? Hemos visto hasta ahora que la relación entre datos en primera persona de la experiencia subjetiva y los datos en tercera persona de la conducta, la identificación psicofísica y el estado funcional, ha sido tratada, en el mejor de los casos de manera reductiva y en el peor de los casos de modo que quedaba simplemente ignorada. Hemos visto que para el dualista una reducción de datos en cualquier término es completamente imposible, ya que la ontología de unos datos y otros son diferentes y están separadas. Para el conductista no es importante averiguar qué ocurre con los datos subjetivos (en primera persona), porque todo lo que podemos saber de los datos objetivos (en tercera persona), expresados en disposiciones, nos dice todo lo que queremos saber sobre los supuestos datos subjetivos. Para el teórico de la identidad, la reducción de los datos subjetivos a datos objetivos es su apuesta principal, pero se ha quedado en *standby* hasta que la investigación descubra los modos de explicar la reducción interteórica, un camino que ha quedado aun más relagado con la aparición de posturas naturalistas antireduccionistas o basadas en correlaciones. Y finalmente para el funcionalista la relación entre un tipo de datos y el otro es posible sólo como instancia de un sistema físico consciente capaz de articular lo subjetivo en términos de representaciones, con lo objetivo proveniente de mecanismos causales.

De acuerdo con el Eliminativismo la subjetividad presente en la “psicología popular” sencillamente serán eliminados de los razonamientos científicos y sustituidos en una nueva teoría neurofisiológica por conceptos más apropiados, referentes a entidades o procesos auténticos. Buscando con ello que los problemas tradicionales de la filosofía de la mente desaparezcan.¹⁹ Lo que el eliminativista afirma es que los estados y proceso intencionales

¹⁹ Cabe lo interesante de resaltar una intuición. El término “eliminativista” proviene de un lugar sospechoso. Alguien que se adscriba al “materialismo eliminativo” no se llamaría a sí mismo eliminativista, porque aquello que estaría “eliminando”, no está siendo eliminado en realidad, simplemente porque desde el principio no existe. Al sano concepto de que todo en el mundo es material lo llaman Eliminativismo, no quienes procuran la “eliminación” de las entidades pertinente, sino quienes sostienen que dichas entidades son algo susceptible de ser eliminado o considerado ontológicamente real, en tanto que son antagonistas de su existencia dentro de la explicación.

de la mente, como las creencias, los deseos, las esperanzas, miedos; qualias, contenidos semánticos, estados intencionales, experiencias subjetivas, o como se le prefiera llamar, a los que aludimos en nuestras descripciones y explicaciones diarias de la vida mental de las personas y sus acciones, en realidad no existen. Estos supuestos estados y procesos son postulados de una teoría profundamente equivocada igual que la que postulaba el flogisto. Alguien inclinado a sostener el Eliminativismo típicamente afirmaría al menos cuatro cosas al respecto de las experiencias subjetivas de psicología popular:

1. “Creencia”, “deseo” y otras expresiones familiares de estados intencionales se encuentran entre los términos de la teoría de la mente de sentido común. Esta teoría es en ocasiones llamada “psicología popular”.
2. La psicología popular es una teoría seriamente equivocada. Muchas de las afirmaciones que hace acerca de estados y procesos que dan lugar a la conducta y muchas de las presuposiciones que hace sobre esas afirmaciones, son falsas.
3. Una ciencia madura que explique cómo trabaja la mente y/o el cerebro y cómo produce el comportamiento públicamente observable, no se remitirá a los estados y procesos intencionales de sentido común que expresa la psicología popular. Las creencias, deseos y demás, no serán parte de la ontología de una psicología científica madura.
4. Los estados intencionales de la psicología de sentido común no existen. (Stich, 1993, pág. 2).

Daniel Dennett es uno de los filósofos que más ampliamente ha defendido un propósito subversivo en contra de los elementos de la psicología popular. Su intento eliminativista va en función de echar por tierra la idea, que es, para la mayoría de las personas, incluyendo, científicos, filósofos y personas sin inclinaciones filosóficas ni científicas, la más obvia de todas las ideas, aquella que tiene que ver con lo que hemos expresado de distintos modos a lo largo de este trabajo, pero que en términos filosóficos integra la noción más básica de la psicología popular: los Qualia. Dennet afirma que la experiencia consciente no tiene ninguna propiedad que sea especial en ninguno de los

sentidos en que se ha supuesto que son. Y propone un proyecto razonable: “cambiar el peso de la prueba, de suerte que quien desee recurrir a propiedades subjetivas y privadas deba mostrar primero que, al proceder así, no está cometiendo un error” (Dennett, 2003, pág. 215).

Para hacerlo Dennett postula lo que da por bien llamar “Generadores de intuición” que permiten a quien investiga descubrir la naturaleza de las propiedades subjetivas y finalmente confrontarla con su eliminación. Pasemos revista a esta estrategia con detenimiento y veamos si logra conducirnos hacia la superación de nuestra propia psicología.

Generador de intuición 1. Mirando a alguien comer coliflor

Mirar a alguien comer algo que no es de nuestro agrado nos hace preguntarnos cómo es posible que a esa persona le guste, y suspicazmente nos convencemos de que es porque debe saberle diferente. Eso es plausible porque para nuestros adentros ha ocurrido algo similar, cuando al probar el primer sorbo de jugo de naranja nos sabe “más dulce” que el segundo sorbo, si entre uno y otro hubo un bocado de panqué con mermelada. El punto central de este generador es mostrar que existen diferencias notables entre sensaciones, de suerte que es correcto hablar de “la forma en que a A le sabe X en el momento t” siendo diferente de “la forma en que a A le sabe X en el momento t’” y diferente también de “la forma en que a B le sabe X en t’”. No obstante de acuerdo con Dennett, en esa forma de hablar ya se está cometiendo el error fundamental de suponer que existe una propiedad residual y que ella debe tomarse en serio para señalar las diferencias. El paralelo del error es obvio, jamás podrá un individuo saber cómo es ser un murciélago o si ve lo mismo que otro individuo distinto, en cuanto a colores, cuando ambos miran un cielo claro y “azul”. La psicología popular nos convence de la realidad de esas propiedades especiales; nos hacen destilar de ellas el resto del mundo para poder definir las.

Generador 2: La máquina catadora de vinos

¿Podrán los catadores de vino humanos ser reemplazados por sistemas computacionales expertos en control de calidad y clasificación? Es posible imaginar un aparato que reemplace las papilas gustativas y los órganos olfativos. Tales dispositivos realizarían las evaluaciones, descripciones y clasificaciones de los vinos incluso mejor que los catadores de vino humanos, porque están programados con precisión, consistencia y aprendizaje acumulativo. Sin embargo, llegamos a pensar que no importa cuán “sensible” y “discernidor” sea el sistema nunca será capaz de disfrutar lo que los humanos disfrutamos cuando catamos o bebemos vino, a saber, la experiencia subjetiva indiscernible. Los términos de la psicología popular tienden a conducir al pensamiento hacia reflexiones de esta naturaleza. Suponer que las características paradigmáticas de los *qualia* escapan no solamente de las pruebas y comparaciones verbales sino de cualquier prueba objetiva fisiológica o conductual, es la intuición popular que Dennett pretende refutar, lo cual constituye en su opinión el error fundamental de la investigación sobre la conciencia. Las concepciones tradicionales de los *qualia*, han supuesto que son en esencia directamente accesibles a la conciencia, que son cualidades fenoménicas inmediatas, sensaciones crudas o demás, que son las cualidades que “parecen interponer un obstáculo inevitable e insuperable para el funcionalismo, o en términos más amplios, para el materialismo o, en términos aún más amplios, para cualquier punto de vista o tratamiento del mundo que sea puramente objetivo y de tercera persona (Nagel una visión de ningún lugar)” Pero en realidad “Los *qualia* ni siquiera son algo sobre lo que nada puede ser dicho; *qualia* es un término filosófico que no fomenta más que confusión, y a final de cuentas no refiere en absoluto a propiedad ni rasgo alguno” (Dennett, Quinear los *Qualia*, 2003) De acuerdo con Dennett todas estas supuestas cualidades son en esencia insensibles para el Eliminativismo, es necesario superarlas y excluirlas de explicaciones auténticas.

De acuerdo con Dennett atribuir a los *qualia* las cuatro características puede resultar demasiado para algunos autores. Shoemaker, por ejemplo, restringe su naturaleza sólo a “los rasgos cualitativos o fenoménicos de las experiencias sensoriales, cuya posesión hace que las experiencias se parezcan o difieran entre sí cualitativamente, de la manera en que lo hacen”. A lo que Shoemaker se refiere con esta afirmación es que lo cualitativo de una

experiencia es su característica intencional, en el modo como hemos analizado. Fenoménico concuerda con que entre las propiedades no intencionales de las experiencias subjetivas está su dinámica/estructura fisiológicas; esta clase de rasgos de las experiencias subjetivas no es accesible a la introspección, así que no es una propiedad fenoménica. Dennett considera “lo fenoménico” (y en distintos lugares lo cualitativo como siendo lo intencional de las experiencias) dentro de los conceptos tendenciosos y nada evidentes que recuerda las formas inefables, privadas y directamente aprehensibles en la conciencia, así que finalmente la discusión permanece en los primeros términos.

Generador 3. El espectro invertido

Ya lo hemos revisado antes pero conviene reanimarlo dentro de esta perspectiva ¿Cómo sé que tú y yo vemos el mismo color subjetivo cuando miramos algo? Ambos aprendimos a reconocerlo a través de nombrarlo, viendo objetos con colores públicos, así que nuestra conducta verbal coincidirá aunque cada cual vea un color distinto. Algunos teóricos están inclinados a pensar que, aun cuando esta intuición es muy sólida, la tecnología podría en principio cerrar la disputa. La teoría de la identidad es especialmente optimista en este punto.

Generador 4. La máquina de estimulación cerebral

Supongamos una tecnología ficción, una máquina neurocientífica ajustada a la cabeza de un individuo cuya función es la de transmitir al cerebro de otro individuo la experiencia visual del primero. Con los ojos vendados el segundo individuo es capaz de informar con suma precisión lo que está viendo, un hermoso cielo amarillo y extensos pastizales rojos. Los qualia de la experiencia del primer individuo definitivamente son diferentes, ha sido empíricamente comprobado. No obstante, cuando se estaba haciendo público el descubrimiento, el técnico operador calibra la máquina, girando 180 grados una perilla y listo, ahora el reporte es “normal”, cielo azul y pastizales verdes. Surge el problema, no hay manera de saber la posición correcta de la perilla, ipso facto no hay manera de realizar una comparación intersubjetiva de los qualia.

Generador 5. La jugarreta neuroquirúrgica

La debilidad del generador anterior es que las experiencias que se comparan están en mentes diferentes. Así que esta es una versión mejorada. Alguien se despierta una mañana para encontrarse con que los pastizales se han vuelto rojos, el cielo amarillo, etc. Nadie más se comporta como si algo raro ocurriera, así que el problema debes ser el sujeto en cuestión. Él concluye naturalmente que está sufriendo una inversión de sus qualia. Durante el transcurso de su historia descubre que unos neurocientíficos malignos son los responsables de su mal. Considerando todo esto resulta que los qualia tienen cierto poder explicativo, y que su presencia es no menos que acertada en una explicación de las causas de un mal razonable, tal como observar alteraciones en la percepción. Los qualia parecen propiedades aceptables, porque las proposiciones acerca de ellos parecen intervenir en las explicaciones acerca de algo más que sólo lo que ocurre “dentro” del sujeto, además intervienen en su conducta, así que puede ser que tengan características causales y así sucesivamente; constituyen a final de cuentas una verificación empírica en tercera persona acerca de la experimentación neurofisiológica detallada de lo que origina el cambio de la percepción. Pero de, acuerdo con Dennett, toda esta explicación en la que los qualia son indispensables es un error y puede ser visto mediante el siguiente generador.

Generador 6. La neurocirugía alternativa

Hay al menos dos formas de explicar que el neurocirujano maligno pudiera alterar los qualia del sujeto:

1. Invertir sus canales productores de qualia, digamos hipotéticamente, en el nervio óptico, de modo que la información que viajaba “de abajo a arriba” hacia el cerebro ahora viaja “de arriba a abajo”, ex hipótesis esto invierte los qualia.

2. Invertir los accesos a la memoria capaces de comparar los tonos actuales con los del pasado, ex hipótesis esto no invierte los qualia, sólo las disposiciones de actuar dependientes de la memoria. Resulta tentador suponer que la evidencia neurofisiológica tendría cabida en la explicación.

Vemos en este generador que el sujeto que se da cuenta de la inversión de sus qualia no lo hace por una alteración real de sus qualia; el ejemplo propone la posibilidad de llegar a la misma conclusión de qualia alterados por una vía que no sea la de la alteración de sus qualia. *Ex hipotesis*, las dos intervenciones quirúrgicas causan los mismos efectos introspectivos, pero sólo una de ellas altera los qualia, así que no hay nada en la experiencia del sujeto que favorezca una hipótesis sobre otra. De modo que los estados de sus propios qualia deben resultarle tan incognoscibles como el estado de los qualia de cualquier otra persona. Si existen los qualia, éstos son aún menos accesibles a nuestra comprensión de lo que habíamos creído. No sólo no son posibles las comparaciones intersubjetivas, sino que tampoco podemos saber en nuestro propio caso si nuestros qualia han sido invertidos, al menos no por introspección. Si estamos equivocados sobre nuestros qualia, es un tema que debe explorarse con detalle, con ejemplos menos de ciencia ficción, más apegados a eventos que realmente pueden ocurrir.

Generador de intuición 7: Chase y Sanborn

Nos cuenta la historia de dos catadores de café trabajadores de Maxwell House, Chase y Sanborn que comparten un problema: a ambos solía gustarles el café de Maxwell House y ahora a ninguno le agrada. Chase afirma que ya no disfruta ser catador de café, antes pensaba que el café que producía Maxwell House era el más sabroso del mundo, y aunque el sabor sigue siendo exactamente el mismo que cuando llegó, él se ha vuelto un bebedor más sofisticado; sus gustos han cambiado. Por otro lado Sanborn afirma que le ha pasado algo similar, aunque sus gustos no han cambiado, algo ocurrió con sus “degustadores”; sus papilas gustativas se modificaron. El café no le sabe igual que antes, si así fuera le seguiría gustando, pero sigue pensando que es el mejor café del mundo y que todos los catadores han hecho tan bien su trabajo que no detecta ningún cambio en el sabor, así que debe ser un problema exclusivamente suyo. De modo que el café de Maxwell House a Chase le sabe igual y a Sanborn no, pero a ninguno de los dos le agrada su sabor como en el pasado. ¿Podría alguno de los dos estar equivocado? ¿Su aparente desacuerdo y dificultades con el café podría consistir más en un modo de expresarse que en una diferencia psicológica? Hay tres opciones de lo que puede estarles pasando a cada uno. En el caso de Chase:

a) Sus qualia asociados a la degustación del café han permanecido constantes, mientras que han variado sus actitudes reactivas hacia esos qualia, lo cual ha influido en sus disposiciones de juicio estético.

b) Simplemente se equivoca con respecto a la constancia de sus qualia; éstos se han alterado gradual e imperceptiblemente a través de los años, mientras que sus criterios gustativos no han cambiado.

c) Se encuentra en una situación intermedia en la que tanto sus qualia se han alterado un poco, como sus estándares de juicio estético.

En el caso de Sanborn:

a) Sus qualia se han alterado debido a un trastorno en su estructura perceptual, pero sus estándares se han mantenido constantes.

b) Sus estándares se alteraron sin que se diera cuenta. Por ello evoca un efecto de nostalgia, paralelo al caso de la conocida experiencia de reencontrarte con algún objeto de la infancia y darte cuenta de que es mucho más pequeño de lo que recordabas.

c) Se encuentra en una situación intermedia en la que tanto sus qualia como sus juicios se han alterado un poco.

Dennett considera que al momento de escribir su artículo todos los que escriben sobre qualia concordarán en que se dan todas estas posibilidades para Chase y Sanborn, pero que nadie se sentiría tentado a defender que la opción a) es la verdad en ambos casos, la infalibilidad de las cuestiones privadas y subjetivas. Desde una perspectiva infalibilista los qualia son propiedades de la propia experiencia que no pueden estar sujetas a error a la hora de describirse. Pero qué tal si cambiamos esa perspectiva por una que nos permita colocar el énfasis en que los qualia fueran construcciones lógicas a partir de los juicios de los sujetos sobre ellos: la experiencia de un sujeto tiene el quale F si y sólo si el sujeto juzga que su experiencia tiene el quale F. Tales juicios dan origen a los qualia, según Dennett, con la misma autoridad con la que los novelistas decretan el color de cabello de sus

personajes. Quienes defienden la autenticidad de los qualia desean que la existencia de todos los qualia particulares sea un hecho empírico que se sostenga por sí mismo, si es así, los qualia continúan siendo un reto para el materialismo o la ciencia objetiva en tercera persona.

Generador de intuición 8: La gradual recuperación postoperatoria

Hemos “invertido quirúrgicamente” las conexiones de las papilas gustativas de Chase en la forma imaginaria estándar, después de la operación el azúcar le sabe salada, la sal le sabe amarga, etc. Sin embargo Chase ha compensado posteriormente esta inversión, según lo demostrarían sus comportamientos; ahora dice que el azúcar le sabe dulce y ya no pide sal para su helado; en otras palabras su conducta y sus reportes verbales son ahora indistinguibles de los de los sujetos “normales”. Puede que el ajuste compensatorio se haya dado en un fase temprana que Dennet llama “prequalia”, en la que sus qualia han quedado exactamente como eran antes de la cirugía o que haya sido en una fase posterior o “posqualia” en la que los qualia de Chase no se han normalizado *aunque él crea que sí*, en ambos casos los hechos fisiológicos, por sí solos, no arrojarán ninguna evidencia con respecto al lugar en el que deba distinguirse que los qualia aparecen como propiedades únicas. El supuesto es desde luego que tales propiedades son inmediatamente accesibles, pero ninguna descripción dentro del curso fisiológico nos servirá para localizar la fase correcta de modificación suponiendo que la hubiere. Si como en el generador de intuición 6 el efecto compensatorio se produjo mediante un ajuste del proceso de acceso a la memoria, seguirían siendo posibles dos opciones:

1. Los qualia actuales de Chase son anormales pero con la revisión de su memoria ha ajustados sus recuerdos sobre cómo deben ser las cosas, así que ya no nota anomalía.
2. El paso de la comparación con los recuerdos se produce justo antes que la fase de los qualia, gracias a la revisión, ahora produce los mismo qualia del pasado ante los mismos estímulos.

Estas son hipótesis sustancialmente distintas, pero ninguna de ellas sigue sin ofrecer evidencia, de en qué lugar debemos considerar a los qualia, en el proceso de acceso a la

memoria. La evidencia introspectiva de Chase tampoco resolverá el dilema entre 1 y 2 ya que ex hipótesis, él no podría establecer una distinción comparativa entre ambos casos. El ejemplo de Chase y Sanborn advierte no sólo que los métodos empíricos no lograrían distinguir afirmaciones acerca de los qualia ni su lugar en los elementos de evidencia para explicar experiencias subjetivas, sino que deriva, además, en la pérdida de una de sus características esenciales: lejos de ser propiedades directa o inmediatamente aprehensibles de nuestra experiencia, son propiedades o bien fuera de nuestra comprensión o bien únicamente inferibles de los exámenes “en tercera persona”. Los conductistas, teóricos de la identidad, funcionalista, neurofisiólogos y otros teóricos objetivos de la tercera persona pueden darle el tratamiento que deseen a esa clase de propiedades para estudiarlas. Sin embargo no son qualia, por la sencilla razón de que la relación epistémica que una persona tiene con ellas es exactamente igual a la relación epistémica que tiene con esas propiedades externas, pero detectables como la temperatura ambiente o el peso. La idea de que uno debería consultar a un experto externo y someterse a complejas pruebas conductuales para confirmar qué qualia uno tuvo, nos desvía mucho de la idea original de que los qualia son propiedades con las que tenemos una familiaridad particularmente íntima. Pensemos en el siguiente experimento y sometámoslo a la intuición de Dennett: Con su mano derecha, específicamente con sus dedos índice y pulgar de la mano derecha, pellízquese en una región bien definida del antebrazo izquierdo..

—¿Ya lo hizo?...

—Adelante, ¡hágalo!

¿Sabe usted qué le ocurrió? -Se lo voy a decir- dice el experto: primero, la presión simultánea y combinada de sus dedos índice y pulgar desencadenaron una serie de disparos neuronales que empezaron en los receptores sensores de la región bien definida de su antebrazo izquierdo que usted pellizcó hace unos momentos, aquellos disparos fueron a parar a su espina dorsal y recorriéndola hacia arriba, llegaron a la región conocida como el tracto de Lissauer, para ingresar luego en el tálamo y en otras regiones basales del cerebro. Entonces la señal corrió hacia la corteza somato-sensorial. Unos pocos centenares de milisegundos después de que usted pellizcara la región bien definida de su antebrazo izquierdo, ocurrió una segunda cosa, una que usted, conoce ya sin necesidad de asesoramiento profesional: ¡usted sintió dolor!

Generador de intuición 9: el bebedor de cerveza experimentado.

Se sabe comúnmente que el disfrutar el sabor de la cerveza es un gusto adquirido. La sabiduría popular afirma que a nadie podría gustarle el sabor del primer trago:

La cerveza le sabe diferente al bebedor de cerveza experimentado. Si la cerveza siguiera sabiéndome como el primer sorbo que probé ¡nunca habría vuelto a tomar cerveza! O para decirlo de otro modo, si mi primer sorbo de cerveza me hubiera sabido como me supo el más reciente, por principio de cuenta ¡nunca habría tenido que adquirir el gusto! Me habría encantado tanto el primer trago como el que acabo de disfrutar (Dennett, 2003, pág. 239).

Pero este es un discurso erróneo, nadie llega a disfrutar la manera en que le supo el primer sorbo. Por el contrario, beber cerveza durante un tiempo hace que las personas experimenten un sabor que disfrutan, pero precisamente el hecho de que disfruten el sabor garantiza que no es el que experimentaron la primera vez. Si esto es correcto la concepción filosófica tradicional de los qualia se enfrenta con algunos problemas extras. Si se admite que las actitudes o comportamientos que se tienen ante las experiencias son constitutivas de sus cualidades subjetiva, de manera que al cambiar la reactividad cambie también la cualidad, entonces esas cualidades o experiencias subjetivas dejan de ser intrínsecas y de hecho se convertirían en cualidades paradigmáticamente extrínsecas y relacionales.

Generador de intuición 10. El experimento mundial de eugenesia.

Este hipotético caso nos habla de la feniltiourea, un compuesto orgánico que para tres cuartas partes de la humanidad tiene un sabor muy amargo pero para el resto de la humanidad es tan insípido como el agua, dependiendo del genoma del degustador. Dado que la reactividad a la sustancia se transmite genéticamente, se podría volver paradigmáticamente amarga realizando un experimento genético a gran escala; impidiendo que las personas a quienes les resulta insípida se reproduzcan; así en unos cuantos años la feniltiourea comenzaría a tener la cualidad de ser amarga sin que nadie objetara nada. O realizar el experimento opuesto y convertir la feniltiourea en una sustancia insípida sin siquiera tocarla. Así que ¿Qué concepto de qualia se debe elegir? Los qualia de Chase se han modificado y debe al fin declarar que esa es la causa de que el café antes solía saberle

bien y ahora le sabe mal; o debe decir que realmente sí sabe como siempre una vez que, desde luego, se sustrae el hecho de que ahora sabe muy mal. Al reconocer que la respuesta no es obvia y se objeta que se debe realizar una elección forzosa que separe dos aspectos que se habían mantenido unidos entonces no queda más que apoyar el hecho de que no existe un fundamento firme en la psicología popular para un concepto de qualia. De acuerdo con Dennett, normalmente pensamos de manera confusa y potencialmente incoherente cuando reflexionamos acerca de las manera en que las cosas nos parecen; cuando experimentamos una reactividad ante un estímulo en diferentes momentos los estados pasados y actuales son algo diferente, eso es algo que podemos reconocer, no obstante no contamos con recursos “inmediatos” o directamente aprehensibles en nuestra conciencia para realizar una distinción fina de ellos. De acuerdo con Dennett esto sugiere que los qualia no son esenciales al vocabulario profesional del fenomenólogo o del fisiólogo; una razón más para la tesis central del eliminativismo.

Generador de intuición 11. La cura de la coliflor.

Alguien le promete a Dennet que después de tomar una píldora la coliflor que no le gusta, le sabrá exactamente igual que siempre, pero su sabor le agradará. Él toma la píldora y ésta funciona, pero no hay evidencia en la experiencia de Dennet que le haga saber cuál es el efecto que tuvo la píldora sobre él, si es por el cambio en sus qualia o por el ajuste en sus recuerdos. Aun sin saberlo, no hay razón para suponer que sus experiencias con la coliflor posean ciertas propiedades subjetivas intrínsecas detrás. Pero sea como deba ser, en principio tiene que haber una respuesta correcta a la cuestión de cómo es intrínsecamente para ti ahora, aunque no se pueda expresar.

Generador de intuición 12. La inversión del campo visual creada por el uso de anteojos inversores

Después de usar un par de anteojos inversores durante varios días, los sujetos del experimento se adaptan con un éxito sorprendente. Supongamos que los confrontamos con la siguiente pregunta: ¿Tu adaptación consiste en que has reinvertido tu campo visual o en que has trastocado por completo el resto de tu mente en un sinnúmero de compensaciones y

ajustes? Debe haber una respuesta correcta; el experimento nos recuerda el viejo experimento del espectro invertido ¿Cómo sé que algunas personas ven las cosas al revés (pero están perfectamente acostumbradas a ello), mientras que otras las ven al derecho? Una persona no puede saber si están intrínsecamente al derecho o al revés independientemente de sus disposiciones a reaccionar ante ellas y si no hay algo que nos obligue a tratarlas como intrínsecas, éstas no tendrían, ex hipótesis, ningún papel que desempeñar en una teoría fisiológica. El argumento es que cualquiera puede experimentar la inversión del espectro, desde una perspectiva subjetiva, hay casos reales del mismo efecto como el desplazamiento de sabor en el caso del jugo de naranja y el panqué con mermelada o el efecto del agua caliente sobre el agua tibia que nos hace sentirla fría. Por tanto la inversión es posible. Y por el mismo hecho deben existir los qualia. De acuerdo con Dennett este es un argumento falaz, ya los términos de desplazamiento o inversión del espectro o del sabor, conllevan las implicaciones problemáticas que aquí se exponen.

No basta ocultar nuestras alianzas teóricas hasta el soleado día en que los filósofos completen la difícil tarea de purificar el concepto cotidiano de qualia. A menos que tomemos medidas diligentes para deshacernos de este concepto original y sustituirlo por mejores ideas, seguirá debilitando nuestra imaginación y distorsionando sistemáticamente nuestros intentos por comprender los fenómenos ya vistos. (Dennett, 2003, pág. 244)

Entonces si los qualia son una pésima idea, ¿por qué han parecido ser una idea tan buena? ¿Por qué parece que existen esas propiedades intrínsecas, infabables, privadas y “cualitativas” en nuestra experiencia? Un análisis de sus presuntas propiedades de segundo orden de las propiedades de nuestras experiencias conscientes nos permitirá diagnosticar su atractivo y encontrar sustitutos adecuados.

Lo intrínseco

La definición de un neófito sobre la pelota de Tomás

Esa pelota tiene muchas propiedades, como la de estar hecha de hule de la India, de pertenecer a Tomás, de haber pasado la última semana en el ropero y de ser roja. Todas,

excepto la última, son propiedades claramente relacionales o extrínsecas de la pelota. Pero su color es una propiedad intrínseca.

La definición relacional

El rojo –el rojo público– es una propiedad esencialmente relacional como lo muestran muchos experimentos mentales acerca de las “cualidades secundarias”, recuérdese el ejemplo de la feniltiourea. Al saber que el rojo es público (del mismo modo que el amargor es público), sabemos que es una propiedad relacional, la resistencia a aceptarlo, de acuerdo con Dennett, consiste en aferrarse a que posee una cualidad de intrinsecalidad. ¿Qué otra cosa serviría para distinguir el efecto del color o del sabor que no sea una propiedad intrínseca? ¿Otra propiedad relacional o extrínseca? No servirá de nada decir que es obvio que algunas propiedades de la percepción son intrínsecas, porque pudo parecerle obvio a algunas personas (las consideraciones sobre el dilema de Chase de la imposibilidad de una comparación muestran las dificultades de que sea obvio). Por lo tanto, en términos de proponer una teoría coherente sobre la conciencia, Dennet sugiere suspender las consideraciones sobre la intrinsecalidad de las experiencias subjetivas al menos hasta que se resuelva su definición y su significado, ya que no es posible afirmar esa cualidad de manera clara y coherente en una teoría sobre la conciencia y las propiedades que posee.

La inefabilidad.

Generador de intuición 13. El canto del cenizote.

Jamás he oído el canto del cenizote, ni siquiera grabado, pero tengo la idea general después de leer algunos libros sobre aves y algunos poemas, de a qué se debe prestar atención “una serie de silbidos cortos y agudos, como chirridos: chip, chip o chuc chuc, entre otras 400 formas”. La descripción verbal pone un límite parcial en la posibilidad de las voces, se pueden descartar millones de ellos, pero se mantiene una amplia gama de posibilidades. Un día identifiqué un cenizote e inmediatamente escucho su canto: “Así que así es como suena” me advierto, ostentando un estado mental que alberga presumiblemente

qualia intrínsecos e inefables. Nombro a ese estado “S”, lo cotejo con las descripciones de los libros y observo que las descripciones son verdaderas, exactas y poéticas, pero aun no logran captar el estado de qualia que he llamado S. De hecho por eso necesito el neologismo “S”, mi experiencia perceptual ha identificado una posibilidad lógica que la descripción verbal no ha podido.

Sin embargo, una única experiencia de este tipo no me permitirá saber qué variaciones o constancias físicas en estímulos producirá en mí esa u otra experiencia indistinguible, tampoco sabría si es la misma reactividad en caso de que se presentara algo que fuera desde el punto de vista de todas las medidas físicas, una segunda estimulación idéntica a la primera. En otras palabras cuando oigo por primera vez el canto del cenizote, quizá identifico un detector de propiedades en mí, pero no tengo todavía una idea de cuál es la propiedad que detecta mi recién descubierto detector de propiedades. Podría parecer que ese descubrimiento no agrega nada de conocimiento, pero no es así, quizá no sé describir qué propiedad es o relacionarla con algún rasgo público, pero en un grado mínimo puedo referir a la propiedad que detecté en ese suceso; mi experiencia del canto del cenizote me ha brindado una nueva forma de pensar acerca de sus 400 voces, que es públicamente contrastable.

Generador de intuición 14: la caja de gelatina Jello. Información en el sentido formal de teoría de la información

Un sistema de contraseñas pares es un método de seguridad muy exitoso y lo es porque cuando partes algo en dos, por ejemplo una hoja de papel, se produce un borde de tal complejidad informativa que sería casi imposible reproducirlo mediante una construcción deliberada, el borde se convierte en un mecanismo único de reconocimiento de patrones M, que es un mecanismo instanciado únicamente por su contrapartida. El sistema de contraseña funciona tan bien porque no podemos sustituir nuestro predicado M por una descripción más extensa y compleja; si pudiéramos hacerlo seríamos capaces de generar un algoritmo capaz de producir otra instancia de contraparte de M. Y de acuerdo con Dennett, eso es justamente lo que hacemos cuando parecemos ostentar, “con el dedo mental de la intención interna”, un quale o un estado mental de qualia en nuestra experiencia. Nos referimos a una propiedad –una propiedad pública de fronteras indefinidas- haciendo

referencia a nuestra capacidad personal para responder a ella. Cuando pregunto si tu azul es mi azul o tu do mayor es mi do mayor, lo que se está preguntando es si nuestros perfiles discriminatorios en determinadas variaciones serán más o menos los mismos. Y puede que no lo sean, las personas experimentan el mundo de modo muy diverso y complejo. La experiencia es cambiante.

Generador de intuición 15: la cuerda de la guitarra

Puntea la cuerda mi baja o grave sola, y escucha el sonido atentamente ¿Tiene partes descriptibles o se trata de uno solo e inefablemente guitarreo? Muchos optarán por decir lo segundo. Vuelve ahora a puntear la cuerda sola y lleva con cuidado un dedo al octavo traste, presionándolo ligeramente para crear un sonido “armónico” agudo. De pronto se escucha un sonido ‘nuevo’: de alguna manera “más puro” y, desde luego, una octava más alto. Algunas personas insistirán en que se trata de un sonido completamente nuevo, mientras que otras describirán la experiencia diciendo que “la base se separó de la nota”, dejando sólo la parte alta. Pero luego, con un tercer punteado, se puede oír, con sorprendente nitidez, el sobretono armónico que se aisló la segunda vez que se punteó la cuerda. La homogeneidad y la inefabilidad de la primera experiencia han desaparecido, y han sido reemplazadas por una dualidad tan “directamente aprehensible” y claramente descriptible como la de cualquier cuerda (Dennett, 2003, pág. 257).

De acuerdo con Dennett, la diferencia en la experiencia es asombrosa, pero la complejidad aprehendida en el tercer punteado, estuvo siempre ahí. En otras palabras, aunque la experiencia subjetiva ha cambiado drásticamente, uno sigue respondiendo, como antes a una propiedad compleja tan altamente informativa que prácticamente desafía toda descripción verbal.

Así, cuando le echamos una última mirada a nuestra caracterización original de los qualia como propiedades inefables, intrínsecas, privadas y directamente aprehensibles de la experiencia, descubrimos que no hay nada que llene esos requisitos. En su lugar encontramos propiedades públicas relativa o prácticamente inefables a las que podemos referirnos de manera indirecta por referencia a nuestros detectores de propiedades privadas –sólo en el sentido de que son idiosincrásicas. Y en la medida en que pretendamos aferrarnos a nuestra autoridad subjetiva sobre la presencia, dentro de nosotros, de estados de cierto tipo o con ciertas propiedades, podemos tener alguna autoridad –no infalibilidad o incorregibilidad, sino algo mejor que mera adivinanza-, pero sólo si nos restringimos a propiedades relacionales, extrínsecas, como el poder de ciertos estados internos nuestros de provocar actos de aparente reidentificación. Así, contrario a lo que a primera vista parece obvio, simplemente no hay qualia (Dennett, 2003, pág. 258).

Retrocedamos un poco y recordemos cómo comenzamos. Partimos del hecho de que todo lo real posee propiedades y asumimos que la experiencia consciente es real, así que por transitividad ella debe poseer propiedades. Asumíamos al principio que tales propiedades son las maneras en las que nos aparecen las cosas a los sentidos, las sensaciones, percepciones y otras actividades mentales intencionales. Sin embargo el esfuerzo eliminativista propone que tales propiedades pueden ser tan distintas de las que asumimos popularmente que sería un error darle a cualquiera de ellas el nombre de qualia o de simplemente experiencia subjetiva. El esfuerzo de Dennett se centra en conducir a la reflexión a considerar no sólo que los diversos conceptos de qualia están equivocados, sino que toda la noción preteórica de la ontología de tales propiedades es el error principal y debemos eliminarla de discusiones serias acerca de la conciencia, porque las propiedades que hacen a la experiencia consciente ser lo que es, no son los qualia, son otra cosa. Es conveniente, renunciar a las cualidades de la psicología popular en virtud de construir una explicación sólida de la ontología de la conciencia. Los qualia no ofrecen ninguna garantía en tal explicación.

Esta serie de afirmaciones han sido objeto de no pocas resistencias en el plano teórico e imaginativo del conocimiento popular. A pesar de que la tesis eliminativista está muy presente en la literatura filosófica reciente, es por otro lado, muy difícil encontrarse con que una persona sin inclinaciones filosóficas considere algo como una ‘teoría’ cuando usa cotidianamente términos intencionales para explicar la conducta de otras personas; y mucho menos que piense que está utilizando una teoría cuando reporta sus propias experiencias y deseos. En pocas palabras es muy difícil imaginar una teoría más radical y provocativa para el sentido común que el Eliminativismo. Jerry Fodor ha comentado que si el Eliminativismo es correcto bien podría ser “la más grande catástrofe intelectual en la historia de nuestra especie” (Fodor, 1987, pág. xii) y Stephen Stich nos propone un par de ejemplos para señalar en qué consistiría esta catástrofe.

Comencemos con la historia. ¿Firmó Lincoln la Proclamación de Emancipación porque quería la abolición de la esclavitud? ¿O fue porque pensó que sería un movimiento estratégicamente útil para ayudar a debilitar a la Confederación? Si el Eliminativismo es correcto, entonces ninguna de estas explicaciones tiene posibilidad de ser correcta, ya que no hay ni deseos ni pensamientos. Consideramos la epistemología: Desde Descartes hasta el presente, los epistemólogos han tratado de construir una teoría sistemática que nos diga cuáles de las creencias de una persona

están justificadas y cuáles no. Pero si el Eliminativismo es correcto, no hay creencias justificadas; no hay creencias en lo absoluto. O considerando la antropología: algunos investigadores han pretendido que una variedad de las emociones humanas, como el miedo, la sorpresa y el disgusto, son universales culturales arraigados en la biología; otros han argumentado que todas las emociones son “construcciones sociales”. Pero si el Eliminativismo es cierto, entonces hay algo profundamente equivocado en esta disputa. El miedo, la sorpresa y el disgusto son estados intencionales y el Eliminativismo afirma que no existen tales cosas. Finalmente consideremos la psicología: si el Eliminativismo es cierto, entonces mucho de lo que sucede en la psicología clínica está destinado a ser inútil. Los problemas de las personas no serán remediados mediante la remoción de las creencias irracionales o haciéndoles darse cuenta de sus deseos subconscientes; no hay tales cosas. Y, obviamente, si el Eliminativismo es correcto, entonces como Fodor insiste, muchos psicólogos cognitivos deberían “hacer [su] ciencia de otra manera” O al menos, “deberían dejar de gastar el dinero de los contribuyentes” (Stich, 1993, pág. 1).

A nuestro modo de ver lo provocativo y aparentemente anti-intuitivo de la teoría consiste en el esfuerzo por explicar el fenómeno de lo mental en términos naturales; después de todo qué teoría tendría que ser tomada con seriedad si no cumpliera con la sustitución de aquello que derroca. Concédase el siguiente ejemplo para introducir la disputa. La definición del color de las dos siguientes formas:

O bien:

1) X es rojo si y sólo si X refleja la luz de longitud de onda N, donde N es algún número.

O bien:

2) X es rojo si y sólo si X se ve rojo por los percibidores normales en circunstancias normales.

Una persona sin inclinaciones filosóficas ni científicas podría preferir 2) porque no va más allá de los que sabe cuando comprende el concepto del color rojo. En cuanto accede a la información de rojo, comprende que las cosas rojas son vistas rojas por las personas normales en circunstancias normales, y que las cosas que se ven rojas en esas condiciones simplemente son rojas. Esta es una intuición de su propio sentido común, no necesita saber

nada de longitudes de onda para saber algo acerca del color; hay personas que tienen el concepto de rojo, que puede usar el concepto de rojo y nunca, sin embargo, haber oído de longitudes de onda. Pero en lugar de eso 1) nos ofrece la posibilidad de construir la explicación mediante las condiciones necesarias y suficientes de ser rojo y tales condiciones son nomológicas; caracterizan el concepto en términos de leyes naturales. Así que cabría investigar para el caso de los estados mentales ¿qué ventaja tendría preferir una teoría de la mente en términos de 1) que una concepción de su naturaleza en términos de 2).

Una posible respuesta es la que introduce la tesis del Naturalismo Mecanicista. El desafío a la base del Naturalismo, es que necesita una explicación de cómo la mente se ajusta al resto de la naturaleza; y en última instancia, descubrir cómo los estados mentales con contenido intencional de la psicología popular y la experiencia subjetiva, si es que todavía podemos expresar algo de ellos en términos mecánicos, puede ajustarse al resto de la naturaleza; aceptándola como un rasgo natural del mundo. Consideremos tres opciones.

Lo que los filósofos naturalistas tratarían de explicar es que los elementos de la psicología popular tienen características causales, es decir que se ajustan al orden causal del mundo, al estar regidos por la misma clase de leyes con la que otros fenómenos de la naturaleza se rigen. Paul Grice (Grice, 1957), advirtió que cosas como la representación mental o el significado pueden ser conceptos causales. Usemos un ejemplo. En las siguientes oraciones el concepto de significado se usa de modos diferentes:

- a) Una luz roja significa alto
- b) Esas manchas significan sarampión

En el primer caso el hecho de que una luz roja signifique alto, es sólo materia de convención, no hay nada acerca del rojo que lo conecte naturalmente con detenerse; el color morado hubiera servido igual. Pero en el caso de las manchas no es cuestión de convención que signifiquen sarampión. Se distingue que hay algo en las manchas que las conecta naturalmente con el sarampión y a causa de esto se pueden usar como indicador de la enfermedad. Las manchas son signos naturales o representaciones naturales del sarampión. Del mismo modo decimos que “el humo significa fuego” o “esas nubes grises

significan tormenta” y lo que queremos decir es que el humo y las nubes son signos naturales del fuego y la tormenta. A esta peculiaridad del sentido de significado Grice le llamó “significado natural”, el cual no es sino una clase de correlación causal. Así que si quisiéramos hablar de la teoría causal de la psicología popular, ésta opinaría que las correlaciones causales entre los pensamientos, creencias o deseo y las cosas que representan intencionalmente, pueden formar su base natural.

La relación entre X y Y, cuando X es un significado natural de Y, se denomina Indicación Confiable; en general, X indica de manera confiable Y cuando hay un enlace causal confiable entre X y Y. El humo confiablemente indica fuego, las nubes grises confiablemente indican tormenta y las manchas confiablemente indican sarampión, etc. De manera que:

X representa Y si y sólo si X confiablemente indica Y

Aplicado a los elementos de la psicología popular, podemos decir por ejemplo que una creencia representa Y si y sólo si hay una correlación causal entre la creencia y Y. Es dable esperar que los esfuerzos naturalistas pueden explicar los poderes representacionales de los elementos de la psicología popular, en términos de la noción de indicación. La indicación confiable es relativamente fácil de encontrar, siempre que hay esta clase de correlación causal, hay indicación. La segunda opción es concentrarnos en el éxito en la acción; en los efectos en lugar de las causas. Particularmente en los efectos expresados en comportamientos. Según hemos visto lo que hace una persona es causado por aquello que cree, siente, desea, etc. Qué se cree, siente o desee, depende de cómo pensemos que son las cosas. Por ejemplo, si pienso que no he tomado vino puedo desear vino. Que logre tomar un poco de vino dependerá de si el modo como considero que son las cosas es como de hecho son las cosas. Si deseo un poco de vino y creo que hay un poco de vino en la cava, entonces que consiga vino yendo a al cava dependerá de que esta creencia sea correcta. El éxito de satisfacer nuestros deseos es cosa que depende de si los elementos de nuestra psicología representan naturalmente el mundo de manera correcta.

Finalmente consideremos a los elementos de la psicología popular como un resultado natural con algún propósito o función evolutiva. Y esta tesis está además a la base de las

afirmaciones eliminativistas. Esto nos recordará un discurso del capítulo anterior. Lo que existe en el universo, es materia, existen los átomos, las moléculas, las partículas y sus interacciones físicas y químicas. Tenemos también las células de los seres vivos, las neuronas, sus conexiones, sinapsis y sus corrientes eléctricas, como tal existen los cerebros; (no hay composiciones de la materia que hayan hecho emerger una clase diferente de funcionamiento tal como la mente en el sentido dualista). Existe el Universo y la Naturaleza gobernada por leyes físicas. Lo que sabemos hasta ahora de los seres humanos es que somos un conjunto de miles de millones de células de miles de tipos diferentes. La mayoría de las cuales provienen del óvulo y el espermatozoide. Las células que nos integran están vivas y han realizado un largo viaje evolutivo que nos permite analizar que cada una de ellas es, por decirlo en palabras de Daniel Dennett, un microrrobot, con funciones de diversa naturaleza adaptadas a las exigencias de las presiones evolutivas constantes, gracias a la cual han desarrollado las características que han desarrollado. Sabemos paralelamente que el metabolismo, la autorreparación celular, la reproducción y demás se realizan por medios que no son en absoluto misteriosos, que no requieren de fuerzas especiales diferentes de las naturales para efectuarse; no necesitan de nada que sea como un *elan vital*, una intervención fantasmagórica o un alma. Podemos afirmar que los fenómenos mentales desempeñan un papel en el aumento de la supervivencia del organismo y con ello de la especie, como resultado de la adaptación a presiones naturales constantes. En algunos casos, incluso parece haber una conexión entre ciertos deseos y la respuesta de supervivencia que los acompaña; considérese el deseo de agua. Si el organismo humano no consigue agua, no sobrevive por mucho tiempo; tener estados que nos motivan a obtener agua de seguro forma parte del rasgo natural que es el deseo de agua.

Sin embargo no es lo mismo decir que los deseos deben haber tenido algún origen evolutivo a decir que estos contenidos (los que representan) pueden explicarse en términos de tales propósitos. Tim Crane los explica como sigue:

La biología sostiene que la selección natural ha garantizado que estemos en estados cuya función es causar una situación que incrementa nuestra supervivencia. Estos estados son los elementos de la psicología popular y las situaciones son sus contenidos. Así, conseguir agua aumenta nuestra supervivencia, de modo que la presión evolutiva constante ha garantizado que estemos en estados que nos inducen a conseguir agua. El contenido de estos estados sería algo como “tengo agua”, porque

nuestra supervivencia ha sido incrementada cuando estos estados provocan un estado de cosas en que tengo agua (Crane, 2008, pág. 299).

Pero el punto que queremos resaltar aquí no es en particular el de la representación, sino el hecho de que si un ser tiene deseos, entonces ha evolucionado. Pues la teoría nos dice que un deseo es precisamente un estado en el cual la selección natural ha otorgado cierta función biológica: causar el comportamiento que incrementa la supervivencia del organismo. Este es el punto de vista de la Psicología Evolutiva, cuya pretensión más específica y no poco controvertida, es la de que muchas capacidades y facultades mentales pueden explicarse considerándolas como adaptaciones en el sentido biológico evolutivo, en el que una adaptación es un rasgo o capacidad cuya naturaleza puede ser explicada como producto de la selección natural. La psicología popular, lleva consigo elementos de los que hemos tratado a lo largo de este trabajo; experiencia subjetiva, contenido semántico, estados intencionales, qualia, etc. No queremos apuntar que todos ellos son sinónimos entre sí, sino que pertenecen a la clase de cosas que reciben un peculiar tratamiento de parte de los esfuerzos materialistas en mor de expulsarlos de las explicaciones de la conciencia.

El análisis que hemos hecho hasta ahora funciona como medio para conjeturar que la experiencia subjetiva de los pensamientos, percepciones, sentimientos y demás, posee precisamente la característica de exceder el entendimiento en estos términos. Al punto de afirmar que es necesaria para la explicación completa de la conciencia, y que los esfuerzos materialistas por sí solos carecen de los medios para comprender cómo elementos conductuales, físicos o naturales de la composición de los seres humanos conforman o hacen emerger la conciencia. Hasta este punto hemos sostenido una postura “zombista” frente al análisis de las teorías que figuran en este capítulo. Los esfuerzos sobre la conducta, la identificación psicofísica o los modelos computacionales o mecanicistas de la conciencia, nos arrojaban la idea de que por fuerza se dejaba algo importante sin explicar: el contenido semántico, la experiencia subjetiva, qualia, el “cómo es ser”, el “sentir crudo” etc. Hemos explorado la posibilidad de su eliminación, pero aun en esos términos parece que su eliminación también es un esfuerzo que como se dice popularmente “se sale por la tangente” y no propone una resolución final al problema de la presencia y funciones naturales de las experiencias subjetivas. Así que en el siguiente capítulo nos dedicaremos a echar mano de un esfuerzo capaz de articular los esfuerzos materialistas de datos en tercera

persona con los datos en primera persona de la experiencia subjetiva: la Ciencia de la Conciencia.

CONCLUSIONES. CONSTRUIR UNA CIENCIA DE LA CONCIENCIA

La explosión de trabajos científicos acerca de la conciencia que comenzaron con el conductismo, la teoría de la identidad, el funcionalismo y el eliminativismo con su contraparte naturalista, así como en neurociencias cognitivas y psicología, nos hacen pensar que estamos cada vez más cerca de una genuina comprensión científica de la experiencia consciente. ¿Qué es la ciencia de la conciencia? ¿Qué forma debería adoptar esa disciplina? Si la conciencia es un proceso natural y objetivo, como sugieren las posturas materialistas, entonces es necesariamente una de las verdades fundamentales del ser humano, debe poder explicarse en los mismos límites científicos que la gravedad o la transmisión de caracteres hereditarios. La conciencia es la razón de que la vida tenga sentido, si no fuéramos conscientes, sin duda sería mucho menos divertido vivir, sin mencionar lo inútil de todo aquello en lo que invertimos para experimentar sensaciones. No obstante, parece que la conciencia sigue siendo el fenómeno más misterioso del universo. ¿Por qué somos conscientes? ¿Por qué parece que dentro de nuestras cabezas está proyectándose una película multitrack, en 3D con sonido envolvente, aroma, sabor y textura, que nos permite sentir además dolor y hambre, placer y dolor? ¿Por qué no somos sólo robots sintácticos que procesan los estímulos y los transforman en resultados de comportamiento sin experimentar tal película?

Cómo hemos analizado parecen dos fuerzas contrarias las opiniones acerca de estas preguntas, la conclusión es, por un lado que la ciencia por naturaleza es objetiva, mientras que por el otro la conciencia parece, por naturaleza, subjetiva. La tarea de una ciencia de la conciencia sería la de integrar sistemáticamente estas dos clases de datos dentro de un mismo marco de investigación; los datos en tercera persona, o datos acerca de la conducta, procesos cerebrales físicos, funcionales y naturales, y los datos en primera persona, o datos acerca de la experiencia subjetiva. Hemos analizado cómo cuando hablamos de un sistema consciente desde el punto de vista de la tercera persona, se presentan una serie de conductas o disposiciones de ella o fenómenos físicos, funcionales o naturales. Por otro lado cuando observamos un sistema consciente desde el punto de vista de la primera persona, se presentan una serie de fenómenos subjetivos específicos de cada experiencia sensorial.

Ambos datos en tales condiciones, necesitan una explicación cabal de su naturaleza; no se trata únicamente de su función dentro de una explicación del fenómeno de la experiencia consciente.

Consideremos por ejemplo el caso de un proceso musical. Si observamos a alguien escuchando música, los datos relevantes en tercera persona incluyen aquellos concernientes a la naturaleza de los estímulos auditivo, sus efectos en el oído y en la corteza auditiva, las diversas respuestas de comportamiento del sujeto y cualquier reporte verbal que el sujeto emita; todos ellos son datos en tercera persona, datos objetivos, y necesitan o tienen una explicación objetiva también. Sin embargo, no es todo lo que necesita explicación. Cualquiera que haya escuchado música experimenta otras cualidades subjetivas asociadas. Una ciencia de la música que explique los diversos datos objetivos, pero que no explique los datos subjetivos sería una ciencia de la música seriamente incompleta. Una ciencia de la experiencia musical completa debe explicar ambos tipos de fenómenos, preferiblemente dentro de un mismo marco científico

De acuerdo con David Chalmers (Chalmers, 1999) el problema de la explicación de los datos en tercera persona asociados a la conciencia puede identificarse de dos manera distintas, relacionado cada uno con la explicación de la naturaleza y la función de cada tipo de datos. El problema de la explicación de los datos objetivos asociados a la conciencia es en ocasiones llamado el “problema fácil”. El problema de la explicación de los datos en primera persona asociados con la conciencia es llamado en ocasiones el “problema difícil”. Esto no se debe, de acuerdo con Chalmers a que los problemas asociados con los datos objetivos sean en algún sentido triviales, sino porque existen modelos claros para explicarlos. Para explicar los datos objetivos, se requiere explicar el funcionamiento del sistema en términos de la tercera persona. Por ejemplo, para explicar la discriminación perceptual, se requiere explicar cómo un proceso cognitivo puede efectuar la función objetiva de distinguir varios estímulos diferentes y producir respuestas apropiadas; se necesita describir el funcionamiento de un mecanismo que efectúe tal función. En la ciencia de la mente, el mecanismo es usualmente neuronal o computacional. Por ejemplo en el caso de la discriminación perceptual, se especifica el mecanismo neuronal o computacional responsable de distinguir el estímulo relevante; aunque cabe mencionarlo, en muchos casos aún no se sabe con exactitud lo que son los mecanismos, pero en principio no parece existir

un obstáculo para encontrar las explicaciones de manera objetiva. Recordemos el ejemplo de la superveniencia del capítulo 1.

Sin embargo, cuando el mismo tipo de explicación se intenta trasladar al campo de los datos subjetivos el modelo fracasa. La razón es que los datos subjetivos, no son datos concernientes a dinámicas o procesos objetivos. Recordemos el ejemplo de María la neurocientífica propuesto por Jackson; incluso si alguien es consciente o sabe todo al respecto de todas las funciones objetivas en las inmediaciones de la conciencia; cosas como discriminación perceptual, integración, reporte y cosas por el estilo, permanece la pregunta sobre por qué todo ese funcionamiento está acompañado incontrovertiblemente por la experiencia subjetiva. Simplemente explicando las funciones objetivas no respondemos esas preguntas. La lección según Chalmers es que los datos subjetivos son irreducibles a los datos objetivos y viceversa. Es decir, los datos en tercera persona aislados proporcionan un catálogo incompleto de los datos que necesitan explicación: si explicamos solamente los datos en tercera persona no hemos explicado todo. Del mismo modo los datos subjetivos solos están aun más incompletos. Una ciencia de la conciencia satisfactoria debe admitir ambas clases de datos y debe construir una conexión explicativa entre ellos.

Algunas formas de esta conexión son las que proponen algunas de las posturas materialistas del capítulo dos. Una posición intermedia sería la de la Teoría de la Identidad que sostiene que aunque hay dos tipos de datos, podemos explicar los datos subjetivos totalmente en términos del material proporcionado por los datos objetivos. Es decir, explicar totalmente el fenómeno de la experiencia subjetivas en términos de procesos cerebrales. La postura es atractiva, pero como hemos visto existen debilidades que recomiendan permanecer escéptico ante ella. Chalmers argumenta de la siguiente manera en mor del escepticismo ante esta posición:

1. Los datos en tercera persona son datos acerca de estructuras objetivas y dinámicas de sistemas físicos.
2. Las estructuras y dinámicas (de bajo nivel) solamente explican hechos acerca de estructuras y dinámicas (de alto nivel).
3. Explicar las estructuras y dinámicas no es suficiente para explicar los datos subjetivos.

Por lo tanto:

4. Los datos subjetivos no están totalmente explicados en términos de datos en tercera persona.

La premisa 1 habla acerca de que los datos subjetivos siempre se refieren a la dinámica de ciertas estructuras físicas. La premisa 2 dice que la explicación en términos de procesos de este tipo, sólo explican más ampliamente procesos de este tipo. Puede haber grandes diferencias entre procesos, como cuando dinámicas y estructuras de bajo nivel dan lugar a estructuras y dinámicas de alto nivel altamente complejas, pero no hay escape al círculo dinámica/estructura. La premisa 3 encapsula el punto de que explicar estructuras y dinámicas, sólo es explicar funciones objetivas y que explicar tal cosa no es suficiente para explicar datos subjetivos. De estas tres premisas la conclusión se sigue. Lo que no se sigue de este argumento es que los datos subjetivos en primera persona y los datos objetivos en tercera persona no tengan nada que ver los unos con los otros; hemos tratado de mostrar a lo largo de este trabajo que existe una asociación sistemática entre ellos. Existen suficientes razones para creer que la experiencia subjetiva está sistemáticamente correlacionada con procesos cerebrales y con la conducta. Sigue siendo posible que cuando un sujeto tiene un determinado tipo de proceso cerebral, éste viene acompañado de un tipo de experiencia subjetiva asociada. La tarea de una ciencia de la conciencia es, en distintos niveles, distinguir correlación de explicación.

Una ciencia de la conciencia, de acuerdo con Chalmers es enteramente posible, únicamente debemos esperar que esta ciencia adquiera una forma no reductiva. Una ciencia de la conciencia no va a reducir los datos en primera persona a datos en tercera persona, sino que articulará las conexiones sistemáticas entre ellos. Es decir, en donde haya co-variaciones sistemáticas entre dos clases de datos, podemos esperar principios sistemáticos para subrayar y explicar la co-variación. Una teoría de la conciencia será en últimos términos una teoría de estos principios.

Desde luego que esta visión de la conciencia desde un marco científico posee notables riesgos fundacionales controversiales. Se enfrenta a las visiones eliminativas que sostienen que el único fenómeno que necesita explicación es aquel que se refiere al funcionamiento objetivo o que no hay en absoluto datos subjetivos acerca de la conciencia. O a las posturas

conductistas o de identidad que afirman que todos los datos subjetivos son equivalentes a datos objetivos, de procesos cerebrales o conductuales, de modo que explicando estos, explicamos aquellos, o que simplemente aquellos no tienen ninguna relevancia. También hay otras versiones la teoría de la identidad que sostienen que los datos subjetivos podrán ser explicados reductivamente por futuras neurociencia (Churchland 1997) sostiene que las futuras neurociencias podrán ir más allá de las estructuras y dinámicas, en formas que actualmente no podemos imaginar. Finalmente Chalmers sostiene que si podemos encontrar suficientes correlaciones entre estados cerebrales y la conciencia, esto calificará como una explicación reductiva y propone algunos proyectos para hacer posible una ciencia de la conciencia.

Aquí cobra relevancia todo la discusión sostenida hasta aquí. Si lo que hemos articulado hasta este punto es correcto entonces una ciencia de la conciencia debe tomar a los datos subjetivos y los datos objetivos seriamente y debe proceder al estudio de la asociación entre los datos en primera persona y los datos en tercera persona sin intentar una reducción. Y de acuerdo con Chalmers eso es precisamente lo que se encuentra en la práctica. La labor experimental de la mente cuyo material y objeto son las dinámicas y procesos objetivos, que constantemente se encuentra con la pregunta de si los datos subjetivos son reales suele considerarlos importantes y oportunos. Por ejemplo trabajos en psicofísica y psicología perceptual se han preocupado por los datos de experiencias perceptuales subjetivas. En la investigación sobre la percepción inconsciente, la distinción en primera persona entre la presencia y la ausencia de experiencias subjetivas es crucial. Y en años recientes una creciente ola de investigaciones se ha enfocado en las correlaciones entre los datos.

Así que en general, las perspectivas para una ciencia de la conciencia son razonables y esperables. Existen numerosos proyectos experimentales que aunque poseen bases objetivas sobre procesos y dinámicas físicas, en la práctica se encuentran con la intervención de la experiencia subjetiva, y la incorporan a sus hipótesis y contrastaciones. Estos experimentos reconocen los problemas distintivos propios de la explicación de la conciencia y aun así arrojan resultados científicos. Es una pregunta abierta cuán lejos podemos progresar en la correlación pero si seguimos este camino eventualmente podemos alcanzar una teoría de los

principios fundamentales que conectan los procesos físicos con la experiencia subjetiva.
(Chalmers, 2004)

BIBLIOGRAFÍA

- Armstrong, D. M. (1980). The nature of the mind. En C. B. (ed), *The mind/brain identity theory*. Londres: Macmillan. Págs 67 - 79
- Block, N. (1996). What is functionalism. En D. M. Borchert (Ed.), *The Encyclopedia of Philosophy Supplement*. New York: Macmillan.
- Carnap, R. (1988). *La construcción lógica del mundo*. México: Instituto de Investigaciones Filosóficas.
- Chalmers, D. (1999). *La mente consciente: en busca de una teoría fundamental*. Barcelona: Gedisa.
- Chalmers, D. (2004). How can we construct a science of consciousness. (M. Gazzaniga, Ed.) *The Cognitive Neurosciences III*. Massachusetts. MIT Press.
- Churchland, P. (1984/1999). *Materia y conciencia*. Barcelona: Gedisa.
- Churchland, P.S. (1997). The hornswoggle problem. In J. Shear, ed, *Explaining Consciousness: The Hard Problem*. Cambridge, Mass.: MIT Press.
- Corredoira, M. L. (2010). Algunas respuestas a las críticas al materialismo en el problema mente-cerebro. En F. R. C. Diosdado, *Neurofilosofía. Perspectivas contemporáneas*. Sevilla: Thémata/Plaza y Valdés. Págs 129 – 141.
- Crane, T. (1995/2008). *La mente mecánica: Introducción filosófica a mentes, máquinas y representación mental*. México: Fondo de Cultura Económica.
- David, C. (2004). How can we construct a science of consciousness. (M. Gazzaniga, Ed.) *The Cognitive Neurosciences III*.
- Davidson, D. (1970/1981). *Sucesos mentales*. México: IIF-UNAM.
- Dennett, D. (2003). Quinear los Qualia. En M. Ezcurdia, & O. Hansberg (Edits.), *La naturaleza de la experiencia. Vol 1: sensaciones*. México: Instituto de Investigaciones Filosóficas - UNAM. Págs 211 - 262
- Dennett, D. (2006). *Dulces sueños: obstáculos filosóficos par auna ciencia de la conciencia*. Buenos Aires: Katz Editores.
- Descartes, R. (1641/2003). *Meditaciones Metafísicas*. Madrid: Gredos.
- Descartes, R. (1644/1995). *Los principios de la filosofía*. Madrid: Alianza.
- Fodor, J. (1987). *Psychosemantics*. Cambridge: Bradford book/MIT Press.

- Grice, H. (1957). Meaning. *The Philosophical Review*, 377 - 388.
- Jackson, F. (2012). Lo que María no sabía. En M. Ezcurdia, & O. Hansberg, *La naturaleza de la experiencia. Vol 1: Sensaciones*. México: Instituto de Investigaciones Filosóficas.
- Jackson, F. (2012). Qualia epifenoménicos. En M. Ezcurdia, & O. Hansberg, *La Naturaleza de la experiencia. Vol 1: Sensaciones*. México: Instituto de Investigaciones Filosóficas.
- Kim, J. (2001). *Causación epifenomenal y superveniente*. (G. Ollivier, Trad.) México.
- Leibniz, G. (1721/1981). *Monadología*. Oviedo: Pentalfa.
- Lewis, D. (1984). *Un argumento en favor de la teoría de la identidad*. México: Cuadernos Crítica, Instituto de Investigaciones Filosóficas.
- Locke, J. (1690/1999). *Ensayo sobre el entendimiento humano*. México: Fondo de Cultura Económica.
- Nagel, T. (1986/1996). *Una visión de ningún lugar*. México: Fondo de Cultura Económica.
- Nagel, T. (2012). ¿Cómo es ser un murciélago? En M. Ezcurdia, & O. Hansberg, *La naturaleza de la experiencia. Vol 1: Sensaciones*. México: Instituto de Investigaciones Filosóficas.
- Place, U. (1956). Is consiousness a brain process? *British Journal Of Philosophy*, 44 - 50.
- Putnam, H. (1981). *La naturaleza de los estados mentales*. México: Cuadernos de Crítica, Instituto de Investigaciones Filosóficas UNAM.
- Ryle, G. (2005). *El concepto de lo mental*. Barcelona: Paidós Ibérica.
- Sellars, W. (1971). *Ciencia, Realidad y percepción*. Madrid: Tecnos.
- Searle, J. (1985). *Mente, cerebros y ciencia*. Madrid: Cátedra.
- Searle, J. (2002). Twenty-one years in the Chinese Room. (J. P. Bishop, Ed.) *Views into the Chinese Room: new essays on Searle and Artificial Inteligence*, 51-69.
- Searle, J. R. (2000). *El misterio de la conciencia*. México: Paidós.
- Shoemaker, S. (2003). Funcionalismo y Qualia. En *La naturaleza de la experiencia* (M. Ezcurdia, & O. Hansberg, Trads.. México: IIF-UNAM). Págs. 183-212
- Stich, S. (1993). *What is folk psychology*. Piscataway NJ: Rutger University Centre for Cognitive Science.

Wittgenstein, L. (1988). *Investigaciones filosóficas*. México: Instituto de Investigaciones Filosóficas.

Wittgenstein, L. (2007). *Tractatus logico-philosophicus*. Madrid: Tecnos.

ÍNDICE

SUBJETIVO Y OBJETIVO. CONCEPTOS RELEVANTES PARA LA CONSTRUCCIÓN DE LA CIENCIA DE LA CONCIENCIA.....	1
INTRODUCCIÓN.....	2
CAPITULO I.....	7
LA NATURALEZA DE LOS DATOS EN PRIMERA PERSONA	7
I.1. La experiencia subjetiva.....	7
I.2. La perspectiva de la primera persona: una manera de identificar la subjetividad.....	22
I.3. Los datos en primera persona: contenedores de la experiencia subjetiva.....	24
I.4. Irreductibilidad de los datos en primera persona	27
I.5 Posturas filosóficas que apoyan la existencia de la subjetividad de los estados mentales: La inteligencia consciente reside en algo no-físico.....	33
CAPÍTULO II.....	38
LA OBJETIVIDAD DESEADA, MATERIALISMO	38
Y LOS DATOS EN TERCERA PERSONA	38
II.1. Conductismo	38
II.2. Teoría de la Identidad	52
II.2.1. Argumentos a favor de la Teoría de la Identidad.....	60
II.2.2. Argumentos en contra de la Teoría de la Identidad	61
II.3. Funcionalismo.....	62
II.3.1. Argumentos en contra del Funcionalismo	73
II.4. Materialismo Eliminativo y Naturalismo Mecanicista	75
CONCLUSIONES. CONSTRUIR UNA CIENCIA DE LA CONCIENCIA	99
BIBLIOGRAFÍA	105
ÍNDICE.....	108