



UNIVERSIDAD NACIONAL
AUTÓNOMA DE
MÉXICO

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE FILOSOFÍA Y LETRAS
INSTITUTO DE INVESTIGACIONES FILOSÓFICAS

INTUICIONES POPULARES Y CONCIENCIA FENOMÉNICA

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

MAESTRO EN FILOSOFÍA

PRESENTA:

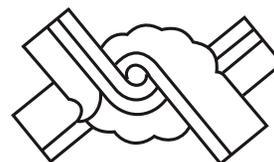
DAVID FAJARDO CHICA

DIRECTORA DE TESIS:

DRA. MAITE EZCURDIA OLAVARRIETA



MÉXICO . D.F. 2011





Universidad Nacional
Autónoma de México



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.

A la memoria de mi padre Álvaro Fajardo

TABLA DE CONTENIDOS

	Página
Tabla de contenidos	i
Agradecimientos	iii
Introducción. Horizonte de la discusión	1
§0.1. ¿Qué es la conciencia? Diversos programas, diversos <i>explanandum</i>	2
§0.2. Requisito de no artificialidad del <i>explanandum</i>	5
§0.3. Sobre el uso de intuiciones	7
§0.4. Argumento central	11
Capítulo 1. Cuatro tesis de la filosofía experimental	14
§1.1. La naturaleza de las intuiciones	14
§1.2. Cuatro tesis de la filosofía experimental	18
§1.2.1. Tesis I	19
§1.2.2. Tesis II	20
§1.2.3. Tesis III	21
§1.2.4. Tesis IV	22
§1.3. Algunos ejemplos	24
§1.3.1. Knobe, J. (2004) - Psicología Social	24
§1.3.2. Machery, E. et. al. (2004) – Psicología Transcultural	27
§1.3.3. Nichols, S. (2004b) – Psicología del Desarrollo	29
§1.4. Alcances de la filosofía experimental	30
§1.4.1. El abanico de problemas es limitado	30
§1.4.2. Tareas disponibles desde la filosofía experimental	32
§1.5. Algunas objeciones al programa de la filosofía experimental	33
§1.5.1. Objeción del experto	33
§1.5.2. Objeción de las intuiciones superficiales	36
§1.6. Conclusiones	38

Capítulo 2. Intuiciones ordinarias en el problema de la conciencia fenoménica	39
§2.1. La enunciación del problema	40
§2.2. Teorías de la mente fisicistas y funcionalistas	41
§2.3. Argumentos de lo concebible	42
§2.3.1. Lo concebible como apelación a intuiciones	43
§2.3.2. Argumento zombi	44
§2.3.3. Argumento de los espectros invertidos	46
§2.4. 'Qualia' como refinamientos de una teoría popular	47
Capítulo 3. ¿Qué ha estado mal en la sistematización de la comprensión ordinaria?	51
§3.1. Knobe & Prinz: Asimetría en atribuciones a entidades colectivas	51
§3.1.1. Dificultad 1: Tales atribuciones no son genuinas	54
§3.1.2. Dificultad 2: Los estados concientes son estados funcionales	55
§3.2. Sytsma & Machery: Percepciones visuales y dolores	56
§3.2.1. Dificultad 1: “Ver rojo” se dice de varias maneras	59
§3.2.2. Dificultad 2: Una premisa de implementación oculta	61
§3.3. Dificultades especiales de este proyecto	62
§3.3.1. Problema con el diseño basado en atribuciones	62
Conclusiones. ¿En dónde quedamos?	66
§4.1. Repaso del argumento	66
§4.2. Consecuencias para el trabajo experimental	68
§4.2.1. Restricción general	68
§4.2.2. Restricción de la intraducibilidad	71
§4.3. Consecuencias para el trabajo teórico	72
Apendices	75
Referencias	79

AGRADECIMIENTOS

En primer lugar, a Maite Ezcurdia quien fue asesora de esta tesis, por la manera rigurosa en la que canalizó mi entusiasmo sobre estos temas. Al Instituto de Investigaciones Filosóficas y a su clima académico tan enriquecedor. A los lectores de esta tesis de quienes recibí provechosos comentarios para mi trabajo, a Ángeles Eraña, Miguel Ángel Fernández, Olbeth Hansberg y Laura Duhau.

Gracias también a quienes siempre estuvieron dispuestos a leer apartes de esta tesis y formular críticas y sugerencias que mejoraron en buena parte la calidad del resultado final, a Santiago Arango, Oscar Fajardo, Rodolfo López, David Rey, Carlos Mario Muñoz y José Tovar. Gracias especiales al profesor Shaun Nichols de la Universidad de Arizona quien fue anfitrión de una visita de verano que realicé en el 2010 para discutir con él algunos puntos centrales de esta tesis.

A mi familia todo el agradecimiento, especialmente a mis dos madres: Amilvia y Sonnia. A Liliana su cariño y a Oscar su interés y entusiasmo. Esta tesis está dedicada a la memoria de mi padre, quien falleció mientras la escribía.

A los grandes amigos que México me permitió conocer, a Sebastián Rudas, Manuela Eusse, Leonardo Ramos, Zach Pirtle e Isabel Carrera.

Nada de esto hubiera sido posible sin el apoyo financiero que el CONACYT brindó a la realización de mis estudios de maestría con su beca para estudiantes de posgrado, agradezco también la financiación que recibí por medio de su programa de Becas Mixtas para la realización de una visita a la Universidad de Arizona en el verano de 2010. De la misma manera, los proyectos “Significado y Comunicación” (DGAPA IN401909) y “Lenguaje y Cognición” (CONACYT U83004) me ofrecieron recursos el tiempo en que terminé de escribir esta tesis.

Finalmente, al Club Dumas por estar ahí.

INTRODUCCIÓN

HORIZONTE DE LA DISCUSIÓN

Esta tesis se enmarca dentro de la pregunta general acerca de cómo establecer aquello que una teoría de la conciencia fenoménica debe explicar. Defenderé que una buena manera de acercarse a la respuesta de esa pregunta es sistematizar la manera en la que ordinariamente se entiende a la conciencia. Por suerte, tal tarea ya ha sido iniciada. Hay trabajos en Filosofía Experimental que han investigado si la comprensión ordinaria de la conciencia corresponde con la caracterización que desde la Filosofía se ha hecho del *explanandum* de “conciencia fenoménica”. Sin embargo, argumentaré que tales estudios sufren de un grave problema en su diseño experimental que no permite resolver el asunto. Así, mi contribución al debate será tanto positiva como negativa: se defenderá la relevancia filosófica de la investigación empírica acerca de la comprensión ordinaria de la conciencia, pero se criticará la manera en la que se ha hecho hasta ahora, estableciendo algunas restricciones y sugiriendo rutas para futuros estudios.

En este capítulo introductorio presentaré el argumento central de la tesis y reseñaré brevemente el contenido de los capítulos siguientes así como su aporte al desarrollo del argumento central. Pero antes haré algunas consideraciones preliminares.

Hay varios proyectos que buscan obtener como resultado una teoría de la conciencia, en el apartado §0.1. diferenciaré el programa que busca una teoría de la conciencia fenoménica de aquellos que también buscan una teoría de la conciencia, pero con diferente *explananda*. En §0.2. presentaré el requisito de no artificialidad en el *explanandum* de una teoría de la conciencia fenoménica. Se defenderá que tal requisito debe ser satisfecho por cualquier teoría que aspire a ser exitosa; la importancia de que este requisito sea aceptado es fundamental para el desarrollo del argumento central de esta tesis, como se verá posteriormente. Aunque tal requisito no ha sido expuesto explícitamente en la literatura, no parece estar muy alejado del espíritu del proyecto general, así que espero no sea muy polémico. En §0.3. haré una breve ejemplificación de la manera en que se han usado intuiciones sobre la conciencia en la Filosofía Analítica, y en particular en la Filosofía de la

Mente, para diversos fines. Parte del interés hacia las intuiciones y su posterior uso como materia prima para la construcción de teorías filosóficas tiene que ver con una motivación similar a la del requisito de no artificialidad. Finalmente, en §0.4 presentaré el argumento central que defenderé a lo largo de la tesis y haré una brevísima presentación del contenido de los cuatro capítulos siguientes.

§0.1. ¿Qué es la conciencia? Diversos programas, diversos *explananda*

La pregunta “¿qué es la conciencia?” puede entenderse de dos modos: como una pregunta por el *explanans*, es decir, sobre cómo se explica la conciencia dentro del marco de la ciencia; o puede entenderse como una pregunta por el *explanandum*, es decir, sobre qué es aquello que una teoría de la conciencia debe explicar. Esta tesis se encuentra en el espíritu de la segunda pregunta. La utilidad de tal pregunta radica en que una aclaración conceptual del *explanandum* es siempre una tarea importante para el establecimiento del *explanans*.

La delimitación del *explanandum* de una teoría de la conciencia generalmente se inicia con una ejemplificación, toda vez que una caracterización más compleja o requiere de términos técnicos o ya presupondrá parte de la teoría que la explica. Así, se ha hablado de la rojez de una rosa cuando se le ve, del dulce sabor de un pastel, de aquello que sucede cuando nos despertamos y deja de suceder cuando nos dormimos o morimos, del uso de la información perceptiva, de los estados de ánimo, de sentirse feliz o triste, de la percatación de los estados internos, de los dolores, las punzadas, las comezones y los orgasmos.

Parte de la polémica entre diversos teóricos de la conciencia se debe a que varios de ellos buscan explicar rasgos distintos de ella, y al tener *explananda* distintos, estarán explicando cosas distintas. Basta una corta ejemplificación de la diversidad de proyectos y, por tanto, de la diversidad de *explananda*.

Conciencia como “estar despierto” [awakefulness]

La conciencia es un sistema de prendido y apagado en el que los organismos que gozan de tal capacidad se percatan (en el sentido más básico del término) de su entorno. Un sistema será conciente o no, pero tal conciencia admitirá niveles, hay grados diferentes de conciencia (Searle

1996, p. 95) donde cada grado implicará más procesos de atención, percepción y memoria. A saber, un paciente en estado de coma no tendría ningún nivel de conciencia. El primer nivel sería el de *conciencia mínima*, en tal estado los pacientes exhiben alguna pista comportamental de actividad mental conciente, por ejemplo respuestas gesturales sí/no o movimientos relacionados con el entorno que no sean producto de reflejos (Giacino 2002). Un nivel superior sería el de los sueños lúcidos donde, según la visión mayoritaria e intuitiva, se tiene cierta conciencia. Por último estaría el nivel de conciencia de los adultos normales. Seguramente, podrían ubicarse niveles intermedios, algunos animales, bebés, enfermos mentales, etc. Sin embargo, gran parte de la atención se centra en el nivel de la *conciencia mínima*. Según este programa, al responder qué es lo mínimo que necesita un sistema para exhibir un comportamiento conciente se estará explicando la conciencia. El trabajo con pacientes con estados mínimos de conciencia busca encontrar los correlatos neuronales (Zeman 2003) dado que se supone que la conciencia se explicará si se explicitan los mecanismos neuronales subyacentes.

Conciencia como acceso a información

En un célebre artículo, Ned Block (1995) señaló que el término corriente “conciencia” confunde dos “conciencias” distintas. Así, podemos entenderlo a la manera de “conciencia fenoménica”, a la que nos referiremos más adelante, y por otro a la manera de “conciencia de acceso”. Tal conciencia de acceso se refiere a la capacidad de un sistema de tener estados internos cuyo contenido pueda ser usado como una premisa en un razonamiento, pueda controlar la acción y pueda servir para el control racional del discurso. (Block 1995, p. 230) De tal modo, si un sistema cuenta con transductores que conviertan la información que recoge del medio en información que pueda procesar y si tiene acceso a esta información procesada para usarla en su desempeño en el mundo, se afirmará que tal sistema es conciente.

Bajo este *explanandum*, se han propuesto varios modelos explicativos, un modelo inicial fue el del *espacio global de trabajo*, que se propuso como un espacio cognitivo similar al de la memoria de trabajo, tal vez de descomposición más lenta, que podría entenderse como un teatro (Baars 1997) sobre el que se proyecta un foco de luz -la atención- y tras bambalinas estaría funcionando el director de escena -un ejecutivo central- encargado de la unión y coordinación de lo que se presenta a la conciencia.

Nuevas metáforas explicativas han llegado. Dennett (2001) propone la de la fama cerebral: alguna información perceptiva (por ej. auditiva o visual) se hace conciente debido a ciertas interacciones entre mecanismos cerebrales que “compiten” por el control del cuerpo, aquellos que momento a momento van ganando la “competición” son aquellos que logran amplificarse estando disponibles de manera global para varios procesos, incluidos -no exclusivamente- la guía de la acción intencional y la categorización perceptiva. Esta disponibilidad general es aquello que experimentamos como conciencia.

Dentro de este programa, se busca una teoría de la conciencia que responda dos preguntas ¿cómo es que cierta información perceptiva logra hacerse disponible y otra no? ¿y luego qué sucede? (Dennett 2001, p. 161). Más allá de las metáforas, el trabajo en neurociencia de Dehaene & Naccache (2001) apunta a la resolución de esas preguntas. La estrategia general consiste en verificar si los sujetos son concientes de cierta información y luego mirar qué sucede en la corteza prefrontal, área donde se supone se hace tal integración perceptiva.

Conciencia como ligadura perceptiva

El problema de la conciencia se explica en términos del problema de la ligadura [*binding*]. Desde esta perspectiva lo problemático de la experiencia conciente es su unidad. Ahora mismo me fijo en la pantalla de la computadora, reconozco la intensidad del brillo, los diversos colores y formas como una sola cosa, aunque sabemos que en nuestro sistema visual hay regiones especialmente sensibles a estos distintos rasgos. Pero además de eso puedo percatarme de la canción de ABBA que suena en el fondo, de la temperatura que hace que opte por usar un abrigo y de lo suave del sillón donde estoy sentado. ¿Cómo si nuestros canales sensoriales son tan distintos, en tipos de información que procesan y estructuras cerebrales implicadas, nuestra experiencia es unificada? Y tal unidad además de ser experimentada, en ocasiones guía nuestras acciones: la sensación de frío en la noche unida a la percepción visual de un cielo nublado y al sonido de truenos a la lejanía, bien pueden llevarme a decidir salir de casa con una chamarra gruesa.

Dentro de este proyecto están las teorías neurobiológicas de la conciencia que proponen Crick (1994) y Llinás (2001). En general, proponen cierta sincronía de los eventos neuronales que explicarían la experiencia unificada. Por ejemplo, en el caso de la conciencia visual las neuronas encargadas de rasgos diferentes de la percepción, como el movimiento, el color y la forma, se

dispararían sincrónicamente a unos 40 disparos por segundo (40 hertzios). De nuevo, se presupone que la explicitación de mecanismos neuronales subyacentes es suficiente explicación.

Conciencia como “conciencia fenoménica”

Aquello que debe explicar una teoría de la conciencia es lo fenomenológico de nuestras experiencias (Chalmers 1995, 1996). Al ver una rosa soy consciente de su color, un rojo profundo que no es uniforme sino que varía ligeramente de pétalo en pétalo, pero también lo soy de su aroma, un refrescante olor que puede generarme algunos recuerdos. Tales experiencias son subjetivas, no puedo compartirlas con nadie y de tal manera también son privadas, incluso parece difícil describirlas completamente de manera verbal ya que siempre que se intente la descripción no podrá agotar su riqueza. El *explanandum* legítimo de una teoría de la conciencia -sostienen quienes defienden este programa- es el “problema duro”, ¿cómo explicar tales rasgos *prima facie* irreducibles en términos acordes con las mejores teorías científicas que poseemos?

Sobre esta diversidad de programas, que dista de ser exhaustiva, valga la pena señalar que no sólo muestra los diversos significados de la palabra “conciencia”, ni los diversos usos que tiene, sino que apuntan a fenómenos mentales distintos, que bien podrían estar superpuestos, pero que en principio podrían estudiarse como diferentes. En esta tesis me centraré en la llamada “conciencia fenoménica” y en especial en la caracterización básica del fenómeno que se ha presupuesto como *explanandum*.

§0.2 Requisito de no artificialidad en el *explanandum*

Al delimitar un *explanandum* esperamos que éste sea la mejor sistematización realizable de los rasgos relevantes del fenómeno y no una construcción artificial. Tal artificialidad puede entenderse como que el *explanandum* esté contaminado de la teoría que lo quiere explicar o que esté alejado del problema que inicialmente produce el interés filosófico (ver §1.2.1.). Éste desiderata motiva el requisito de no artificialidad en el *explanandum* (RNAE). A continuación desarrollaré este requisito que se impone al *explanandum* de una teoría sobre la conciencia.

La discusión sobre el *explanandum* de una teoría de la conciencia será acerca de qué rasgos tiene la conciencia que deben ser explicados, es decir, cuál es la caracterización básica que debe hacerse del fenómeno. Una vez se tenga un *explanandum* definido y común, se podrá evaluar el éxito de una

teoría de la conciencia y el avance en la explicación. Ahora bien, ¿qué evidencia podemos ofrecer para la construcción de tal *explanandum*? Cada vez que se hace referencia a sus diversos rasgos, lo que se hace es consultar la manera como la conciencia *es* para nosotros. Lo aconsejable es usar esa evidencia en la sistematización de la comprensión que tenemos sobre la conciencia para que pueda ser tratada como un *explanandum* operacional.

Ahora bien, los juicios a partir de los cuales se podrá hacer una sistematización tendrán que cumplir ciertos requisitos. Deberán ser juicios más o menos regulares para asegurar que refieran a rasgos del fenómeno y no a particularidades individuales, deberán ser juicios producidos en condiciones de normalidad y cumplir con cualesquiera otros estándares normativos usuales para la elaboración de descripciones psicológicas.

El requisito de no artificialidad en el *explanandum* es deseable para todas nuestras teorías, sean físicas, biológicas, psicológicas o filosóficas. Sin embargo, la manera particular como debe aplicarse varía entre unas y otras en virtud de los objetos estudiados. Cuando estudiamos fenómenos físicos o biológicos, la delimitación del *explanandum* no debe pasar por nuestro entendimiento del sentido común de tales fenómenos pues esto nos podría llevar a tener que explicar como reales rasgos ilusorios. Sin embargo, en el caso de la conciencia este es un insumo importante porque parte importante del fenómeno está estrechamente vinculado con la manera en la que *es* para nosotros. Este vínculo debe quedar plasmado en la caracterización del *explanandum*.

Eso no significa que la investigación ulterior no pueda arrojar como resultado que unos u otros rasgos sí son aparentes o ilusorios, en el *explanans* bien pueden ser explicados, eliminados o reducidos. Pero, aún así, sus lugares en el *explanandum* no se pondría en duda; una teoría debería explicar esos rasgos incluso si los considera ilusorios. Si aceptamos un *explanandum* artificial sobre la conciencia estaríamos explicando otro fenómeno, no aquello que sucede en nuestras mentes concientes y que inicialmente suscitó el interés filosófico. Este resultado indeseable es el que busca evitar el requisito de no artificialidad del *explanandum*.

No quiero decir con esto que no se pueda establecer un *explanandum* de la conciencia de manera teórica, creo que el teórico de la conciencia bien puede sistematizar la comprensión ordinaria que de ella tenemos de tal manera que se pueda delimitar y caracterizar sus rasgos rigurosamente utilizando los términos técnicos que tiene a su disposición. Tampoco quiero decir que aquellas afirmaciones de la comprensión ordinaria que expliquen la conciencia sean útiles como parte de la explicación, es

decir, como parte del *explanans*. Ese no es el papel que deben tener las intuiciones. Por ejemplo, aunque tuviéramos evidencia a favor de que la inmensa mayoría de los sujetos explican su vida mental conciente de manera dualista (Bloom 2004), esto no serviría como evidencia a favor del dualismo. Los sujetos sí pueden estar equivocados acerca de cómo funciona la conciencia, acerca de su naturaleza metafísica y acerca de cómo podemos conocerla, pero difícilmente estarán equivocados acerca de *cómo es*, es decir, cómo luce para ellos. En otras palabras, la explicación bien puede resultar artificiosa y contraintuitiva, pero la delimitación del *explanandum* no puede serlo.

§0.3. Sobre el uso de intuiciones

La motivación que lleva a explicitar el RNAE es compartida por un gran sector de la Filosofía Analítica contemporánea. Queremos que las teorías filosóficas sobre varios objetos de estudio filosófico, sean estos morales, políticos o epistemológicos, no se aparten de aquello que inicialmente fue interesante, es decir, de la manera en la que tales objetos son para nosotros. Una teoría de la Justicia, del Conocimiento o de lo Bueno, que se aleje de lo que intuitivamente consideramos como justo, como conocimiento o como bondadoso, será una teoría extraña que tal vez no resulte satisfactoria.

De este modo, se ha hecho una invocación frecuente de intuiciones en diversas discusiones, en Filosofía Moral, Teorías de la Justicia, Epistemología y por supuesto, Filosofía sobre la Conciencia. Se ha esperado que tales apelaciones a intuiciones funcionen como evidencia para diversos fines: como premisas para construir argumentos, para validar objetos de estudio, para deslegitimar posiciones contrarias y especialmente en forma de experimentos mentales que le han dado un gran dinamismo a la discusión.

A continuación presentaré dos experimentos mentales. El primero tomado de la Epistemología, un célebre caso Gettier (1963) y el segundo más cercano a la discusión que nos compete, el caso de “María la neurobióloga” (Jackson 1982, 2003). El primero apela a intuiciones que buscan desmotivar una teoría, el segundo, además del uso anterior, adiciona elementos para el *explanandum* del fenómeno. No discutiré su valor o éxito; solo se traerán a colación en este apartado para ejemplificar el uso de intuiciones en la metodología de la Filosofía Analítica, la discusión más amplia sobre la naturaleza de las intuiciones no se presentará sino hasta el primer

capítulo y será en el segundo capítulo donde se discuta la relevancia particular de las intuiciones en la construcción del *explanandum* de una teoría de la conciencia fenoménica.

Casos Gettier

Los casos propuestos por Gettier (1963) funcionan como contraejemplos de la teoría del conocimiento tripartita tradicional. Según tal teoría, un caso de conocimiento será una creencia verdadera justificada. Es decir, debe haber un sujeto que tenga una relación de creencia con una proposición, tal proposición debe ser verdadera y además la creencia debe estar apoyada de alguna manera que garantice que no ha sido sólo adquirida por accidente.

Tal caracterización tradicional se presenta como una enumeración de condiciones necesarias y suficientes para que algo pueda ser llamado “conocimiento”. La estrategia de Gettier (1963) fue apelar a la comprensión sobre el conocimiento mostrando ejemplos donde las tres condiciones son satisfechas pero intuitivamente tales casos no son reconocidos como instancias de “conocimiento”.

El caso de Smith, uno de los casos Gettier, es el siguiente: Smith se ha presentado a solicitar un empleo junto con Jones. Smith, además, tiene la creencia: “Jones se quedará con el puesto”. Incluso, tal creencia es justificada, el empleador le dijo a Smith que así sería. Por otra parte, Smith tiene la creencia: “Jones tiene diez pesos en el bolsillo”. Incluso, tal creencia está justificada, Smith logró ver al bolsillo de Jones cuando estaban muy cerca. Así, Smith por transitividad tiene la creencia: “El hombre que tiene diez pesos en el bolsillo se quedará con el puesto”. Finalmente, el puesto no se lo queda Jones, sino que se lo dan a Smith. Sin embargo, sin saberlo Smith también traía diez pesos en su bolsillo. Así que su creencia: “El hombre que tiene diez pesos en el bolsillo se quedará con el puesto”, fue una creencia verdadera y justificada. Pero, señalan quienes proponen este tipo de casos, parece extraño admitir que fuera una instancia de conocimiento.

Esta estrategia es característica del método del análisis conceptual. Se investiga sobre un concepto intentando dar cuenta de él en términos de algunos más básicos y se somete a prueba tal análisis con contraejemplos. Estos contraejemplos, se supone, tienen carácter evidencial, porque apelan a los usos que le damos al concepto de una manera no teórica, tal uso bien podrá llamarse intuitivo. ¿Por qué es importante apelar a tal comprensión no teórica? La motivación aquí es similar a la de proponer el RNAE a una teoría de la conciencia. Es una manera de garantizar que la teoría en

cuestión trate sobre el conocimiento, tal y como se ha comprendido ordinariamente¹, ya que esa fue la preocupación filosófica inicial. Si una teoría del conocimiento no logra pasar el análisis de contraejemplos Gettier, se calificará como contraintuitiva y será desechada. Por tal motivo se han propuesto maneras de salir adelante, en especial, haciendo cláusulas de excepción a la definición tripartita, e.g. sin suerte epistémica (Unger 1968) o sin falsa evidencia (Feldman 1974).

El caso de María la neurobióloga

El experimento mental originalmente propuesto por Frank Jackson (2003, p. 96) busca resaltar intuiciones acerca de la conciencia que vayan en contra del programa fisicista que pretende reducir la mente, y por tanto la conciencia fenoménica, a estados físicos entendidos de un modo amplio. El papel de tales intuiciones es claro para Jackson: “Nada de lo que se pueda decir en el ámbito de lo físico capta el olor de una rosa, por ejemplo. Por lo tanto, el fisicismo es falso. Según nosotros, éste es un argumento perfectamente bueno. Obviamente, no viene al caso cuestionar su validez, y la premisa es, *desde un punto de vista intuitivo, obviamente verdadera para ellos y para mí*” (Jackson 2003, p. 96, las itálicas son mías). Pero Jackson se enfrenta a una dificultad: “Por desgracia para nosotros, hay muchos a quienes la premisa *no les parece intuitivamente obvia*. La tarea consiste entonces en ofrecer un argumento cuyas premisas *sean obvias para todos*, o al menos para el mayor número posible” (Jackson 2003, p. 96, las cursivas son mías). El argumento, a modo de experimento mental, propuesto por Jackson es el siguiente:

Imagínese una brillante neurocientífica de la visión, María, que ha sido forzada a aprender todo lo que ha aprendido encerrada en una habitación en blanco y negro. Sus estudios los ha realizado en libros y monitores a blanco y negro, y de algún modo se ha evitado que ella misma tenga contacto con la coloración de su propia piel. Ahora bien, el estado de la ciencia de la visión que María ha aprendido es completo, es decir, María ha aprendido todo lo que puede aprenderse sobre lo que sucede cuando alguien ve, por ejemplo, un tomate maduro y usa términos como “rojo” para identificar su color. Luego de un tiempo, María logra salir de la habitación y se encuentra por primera vez con un tomate. ¿María aprende algo?

¹ Cabe la posibilidad de que tal comprensión ordinaria no sea homogénea. Evidencia a favor de tal diversidad de intuiciones epistémicas han sido presentadas por Weinberg et. al. (2001). Sus resultados sugieren que sujetos del este de Asia o de la India sub-continental consideran los casos Gettier como instancias de conocimiento. El cómo tratar esta evidencia es asunto de debate en la Epistemología, las dos peores opciones para la epistemología analítica tradicional son, aceptar que la teoría tripartita es sólo una teoría sobre el conocimiento de los Occidentales o aceptar que las intuiciones no serán una fuente de evidencia importante dada su diversidad demográfica.

Al respecto dice Jackson: “Parece perfectamente obvio que aprenderá algo acerca del mundo y de la experiencia visual que tenemos de él [...] es ineludible que su conocimiento previo era incompleto. Sin embargo, tenía *toda* la información física. *Ergo*, hay más conocimiento por obtener que eso, y el fisicalismo es falso” (Jackson 2003 p. 100).

El experimento mental de Jackson persigue dos objetivos. Primero, rescatar intuiciones no fisicistas que funcionen para desmotivar tal programa. Si el experimento mental resulta exitoso, es decir, si la conclusión: “María aprende algo nuevo acerca del rojo del tomate una vez lo ve, a pesar de tener previamente toda la información física que sobre el ver rojo puede tenerse” resulta aceptable, quedaría en evidencia que un programa que busque reducir la conciencia fenoménica a hechos físicos, y por tanto susceptibles de ser estudiados de manera completa por una ciencia física, no tendrá éxito. Segundo, delimitar el espacio conceptual de explicaciones que pueden darse sobre la conciencia, una vez el argumento tenga éxito la naturaleza del fenómeno que se quiere explicar escapará de las explicaciones fisicistas y por eso el abanico de explicaciones se reduce. Si bien el autor elige el epifenomenismo como marco explicativo², las intuiciones antifisicistas a la que apela Jackson bien podrían estar acordes con otras posiciones no fisicistas. Una vez se conceda la argumentación de Jackson se imponen restricciones al *explanans* de la teoría vía la delimitación del *explanandum*: el fenómeno de la conciencia será de tal naturaleza que lo harán irreducible a una teoría física, así una teoría de la conciencia *debe dar cuenta* de tal irreducibilidad.

¿Cómo podemos entender la pretensión de obviedad de Jackson respecto a las premisas de su argumento? Parece que esperar que una premisa sea obvia es lo mismo que esperar que una premisa sea intuitiva. ¿Qué quiere decir esto? Una premisa será intuitiva si resulta *prima facie* aceptable para quien la evalúa. Así, una premisa será intuitiva si se encuentra acorde con la comprensión que sobre el fenómeno se tiene. Ahora bien, si se quiere satisfacer el requisito de no artificialidad del *explanandum* debemos esperar que la comprensión a la que se apele sea la comprensión ordinaria. La comprensión de los teóricos bien podría ser muy refinada o estar permeada de teoría, haciendo que la no artificialidad no esté garantizada.

La referencia usual a las intuiciones ha hecho que algunos filósofos y psicólogos se interesen en hacer una investigación empírica acerca de cuál es la comprensión ordinaria que de la conciencia se

² Según el epifenomenismo la conciencia fenoménica es un residuo psicológico sin ningún poder causal sobre aquello que la produce. Una metáfora que se ha usado con frecuencia para ejemplificar el epifenomenismo es el de la relación entre una locomotora y el vapor que produce. Si bien el funcionamiento de la locomotora produce el vapor, una vez producido, éste no afecta causalmente el funcionamiento de la máquina.

tiene (Arico et. al. por aparecer, Knobe & Prinz 2008, Sytsma por aparecer, Sytsma & Machery 2009, Sytsma & Machery por aparecer, Huebner 2009, DeBrigard por aparecer). Este programa empírico, aunque interesante por mérito propio, resulta filosóficamente provechoso en la aplicación del requisito de no artificialidad del *explanandum*.

§0.4. Argumento central

En esta tesis defenderé a partir del RNAE cómo la comprensión ordinaria sobre la conciencia es determinante para responder qué *explanandum* debemos aceptar como legítimo de una teoría de la conciencia. Para ello, desarrollaré el siguiente argumento en el transcurso del texto:

- (1) Las teorías de la conciencia deben satisfacer el requisito de no artificialidad del *explanandum* (RNAE).
- (2) Buena parte del *explanandum* de las teorías filosóficas de la conciencia se ha delimitado apelando a intuiciones.
- (3) (por 1 y 2) Debemos saber si las intuiciones sobre las que se han construido el *explanandum* de las teorías de la conciencia son de tal tipo que hacen que éste cumpla con dicho requisito.
- (4) La manera para probar si tal requisito se satisface es sistematizando la comprensión ordinaria de la conciencia.
- (5) Tal sistematización no se ha podido lograr por un defecto central en el diseño experimental.
- (6) (por 4 y 5) Aún no podemos determinar qué teorías o argumentos satisfacen el RNAE.
- (7) Así, requerimos de nuevo trabajo experimental que supere el defecto señalado.

Como es de esperarse algunas de las premisas de este argumento resultarán más controversiales que otras, esto se verá reflejado en la manera en la que los capítulos se centrarán en ellas. Espero haber motivado satisfactoriamente la premisa (1) en el apartado anterior acerca del RNAE. El objetivo del capítulo 2 es ofrecer apoyo a la premisa (2), se hará especial énfasis en los argumentos de concebibilidad y en cómo puede entenderse la apelación a intuiciones en ellos. Aceptadas (1) y (2) la premisa (3) será una consecuencia razonable. La premisa (4) estará indirectamente motivada en el capítulo 1. La premisa (5) es el núcleo del capítulo 3. Se discutirán de manera detallada dos estudios y se presentará un argumento en contra de la estrategia que se usó en esos diseños experimentales. La premisa (6) es una conclusión lamentable de las dos anteriores. Finalmente, en el apartado de

conclusiones se establecerán unas restricciones y se sugerirán algunas rutas para la satisfacción de (7).

La anterior fue la relación entre el argumento central y los capítulos de la tesis. A continuación presentaré brevemente el temario de los capítulos. El primer capítulo no tiene una vinculación particular con ninguna de las premisas, pero servirá de marco si cumple dos objetivos. Primero, se hará una presentación general sobre las intuiciones de tal forma que motive el interés por las intuiciones generadas por teorías tácitas. Para ello se discutirán las tesis generales del programa de la Filosofía Experimental y las discusiones metafilosóficas que conlleva: el papel de las intuiciones en la argumentación de la Filosofía Analítica y el compromiso empírico que se adquiere al apelar a intuiciones. También se ejemplificarán algunos trabajos relevantes con el objetivo de ubicar esta tesis como un desarrollo de ciertas ideas filosóficas, aunque también es heredera de ciertos trabajos en Psicología Experimental. El segundo objetivo es brindar una justificación general del enfoque que se ha elegido, mostrando la relevancia general de la Filosofía Experimental se espera ambientar la relevancia particular de la Filosofía Experimental en el caso de la discusión sobre conciencia fenoménica.

El segundo capítulo motivará de dos maneras distintas la entrada de consideraciones empíricas sobre las intuiciones de cara al problema de la conciencia fenoménica. Primero, se mostrará cómo los argumentos de la concebibilidad antifisicistas, en el que juega un papel central la conciencia fenoménica (o *qualia*), pueden entenderse como una apelación a intuiciones. Segundo, se presentará la reconstrucción planteada por Dennett de la delimitación teórica que se ha hecho de los *qualia*, según la cual se ha realizado a partir de intuiciones preteóricas, sin ser él mismo defensor de la “conciencia fenoménica” como *explanandum*.

El tercer capítulo ofrece un recorrido por dos de los intentos de sistematizar la comprensión ordinaria acerca de los estados mentales concientes y responder a la pregunta de si tal comprensión corresponde con el concepto de conciencia fenoménica. Es decir, a la pregunta de si los legos en filosofía y psicología comprenden la conciencia de un modo similar al concepto de conciencia fenoménica. Los experimentos a partir de los cuales se intenta la sistematización tienen similitudes importantes en sus diseños, pero llevan a sus realizadores a conclusiones distintas. Mostraré problemas que subyacen a ambos trabajos, presentaré evidencia a favor de algunas polémicas internas a los experimentos y señalaré la razón por la que no logran resolver la pregunta que nos interesa.

En el cuarto capítulo a modo de conclusión propondré un par de restricciones que futuros diseños experimentales deben tener en cuenta si se desean superar las críticas que he formulado, señalaré una ruta inicial para explorarse en esta dirección y se hará un balance de la discusión.

CAPÍTULO UNO

CUATRO TESIS DE LA FILOSOFÍA EXPERIMENTAL

Este capítulo inicial tiene el objetivo central de servir de marco. De este modo, dejaré el apoyo de las premisas puntuales del argumento central para hacer algo más general. Se discutirá la naturaleza de las intuiciones y se especificarán cuáles serán relevantes para las discusiones filosóficas. Luego se hará una discusión general sobre el programa de la Filosofía Experimental y se defenderá de algunos críticos. Así, en §1.1. discutiré la naturaleza de las intuiciones, presentaré el modelo de Cummins que resulta lo suficientemente general para ser aceptado sin resultados muy polémicos y defenderé la idea según la cual las intuiciones filosóficamente relevantes serán aquellas generadas por teorías tácitas. En §1.2. presentaré las tesis principales que comparten aquellos que hacen trabajo en Filosofía Experimental; en §1.3., a modo de ejemplificación, haré una presentación breve acerca de algunos trabajos fruto de su enfoque; en §1.4. discutiré los alcances y limitaciones del programa, buscando responder a las preguntas sobre qué esperar y qué no del programa de la Filosofía Experimental; y por último en §1.5. revisaré algunas críticas que han surgido alrededor de este programa y mostraré algunas maneras de salir adelante a ellas.

§1.1. La naturaleza de las intuiciones

Iniciaré con una caracterización negativa. Llamaré aquí “intuición” al juicio que no ha sido fruto de un proceso perceptivo, ni de un proceso introspectivo (Sosa 1998). Recuérdese el caso Gettier acerca de Smith (ver §0.3.), estamos dispuestos a negar que es un caso de conocimiento genuino, sin embargo, tal disposición no se debe a algún asunto perceptivo. Por supuesto, hay un proceso perceptivo implicado en el entendimiento del caso, pero la provocación de la intuición no es ella misma fruto de la percepción, el tipo de percepción usada es la misma que se usaría con la lectura de un ejemplo que no provocara tales intuiciones. Tampoco ubicamos el caso fuera de las instancias de conocimiento gracias a un proceso introspectivo, no se hace una revisión exhaustiva acerca de lo que pensamos del conocimiento, no se hace una reflexión en la que recordemos casos de conocimiento y luego las comparemos con el caso de Smith. Incluso, las intuiciones también han

sido calificadas como juicios rápidos o, como en el caso de Smith, juicios reactivos frente a un caso hipotético.

Tampoco serán fruto de un proceso explícito de razonamiento (Shafir 1998; Wisniewski 1998). Es decir, no son resultado de un proceso susceptible de categorizarse dentro de las formas de razonamiento, incluso aunque se acepte que un proceso de razonamiento pueda ser inconsciente. La conclusión, en el caso del experimento mental de María: “María aprende algo nuevo a pesar de tener toda la información física sobre la visión”, no se concluye sólo a partir de las premisas pasadas, no es una inferencia a la mejor explicación ni es una consecuencia necesaria de la historia relatada en virtud de alguna regla inferencial.

Además, calificaré como “popular” a aquella intuición propia de un grupo de sujetos respecto a un fenómeno específico tal que no posean una teoría disciplinar explícita respecto al fenómeno y/o no posean entrenamiento en la disciplina que suele encargarse de fenómenos similares, de tal modo que los haya hecho enfrentarse a problemas y preguntas relacionadas. Así, una “intuición popular” acerca de un fenómeno será *el juicio espontáneo, fruto de un proceso no accesible de manera consciente, compartido por la gran mayoría de los sujetos, normales y competentes del lenguaje, que no posean (o no hayan tenido contacto con) una teoría propia de la disciplina que estudia el fenómeno sobre el que versa el juicio.*

En lo que sigue del apartado se presentará la caracterización de Cummins (1998) de las fuentes de intuiciones. Este acercamiento tiene dos virtudes, por un lado brinda un bosquejo acerca de la psicología de las intuiciones siendo a la vez lo suficientemente general para ser aceptado sin mayores polémicas. El enfoque que se ha privilegiado es el de examinar cuál es la fuente de las intuiciones. ¿De dónde provienen tales juicios espontáneos? La naturaleza de los mecanismos y procesos subyacentes no se tratará, cualquier aproximación sobre ello será neutral respecto a lo que aquí se discutirá.

Cummins (1998) plantea algunos candidatos considerados fuentes de intuiciones dentro de los cuales examinaré las teorías explícitas, las creencias ordinarias y las teorías tácitas. Una *teoría explícita* es aquella teoría que se ha aprendido, generalmente fruto de la educación, acerca de un determinado fenómeno. Si bien el contenido de la teoría es explícito, es decir, el poseedor puede reportarlo, ahondar en él, usarlo para hacer inferencias, etc., no es necesario que el poseedor declare

que ese contenido constituya una “teoría”, y tampoco que el poseedor la identifique con algún nombre.

Las intuiciones que una persona con educación avanzada tiene acerca de la naturaleza de los organismos vivientes son producidas por una teoría explícita que se aprendió en un momento específico y que, seguramente, reemplazó una teoría popular sobre los organismos biológicos. Una teoría explícita, por definición, es una teoría que reemplaza o quizás reivindica una teoría popular, mas no depende en ningún sentido de ella; de este modo, las intuiciones que una teoría explícita podría producir no son intuiciones populares ya que éstas sólo serán compartidas por quienes hayan tenido contacto con dicha teoría.

Hay otra manera en la que una teoría explícita puede generar intuiciones o al menos afectar la producción de ellas. Cummins no advierte tal posibilidad pero resulta importante, así que la mencionaré. A pesar de que una teoría explícita no trate directamente al fenómeno, cabe la posibilidad que sí se puedan hacer ciertas inferencias desde la teoría hasta el fenómeno en virtud de alguna vinculación indirecta. Tómese como ejemplo la posesión del conocimiento explícito acerca de cómo funciona un radio, si bien tal cuerpo de conocimiento no implica de ninguna manera el conocimiento acerca de cómo funciona una computadora, alguien que posea tal teoría explícita tenderá a pensar que el funcionamiento de la computadora se debe a algún mecanismo interno explicable y no a alguna suerte de magia o encantamiento. Alguien que además de no poseer una teoría acerca del funcionamiento de una computadora, no posea una teoría acerca del funcionamiento de un radio o de algún otro aparato, sí podría intuitivamente pensar que la explicación es mágica.

Otra fuente de intuiciones que señala Cummins, son las *creencias ordinarias*, creencias formadas con el paso del tiempo en el seno de un ambiente social y cultural específico y relativo a cada quien, creencias construidas a partir de los libros que leemos, las películas que vemos, las conversaciones que tenemos, etc. Estas creencias ocupan los imaginarios populares, tales como: “Los hombres son más promiscuos que las mujeres”, “Todas las cosas suceden por algo”, “La gente más educada son mejores personas que las menos educadas”.

Las teorías explícitas y las creencias ordinarias son similares en tanto sensibles al contexto y aprendidas. Pero difieren en la sistematicidad y carácter disciplinar de la primera. Es posible que el éxito y la simplicidad de una teoría explícita hayan hecho que ésta penetre la cultura popular, de tal

suerte que existan creencias ordinarias que hayan llegado a una población a partir de la popularización de alguna teoría explícita. Sin embargo, por la justificación, el poder inferencial, y las creencias contextuales que la hacen más poderosa (quien posee una teoría explícita conoce un contexto teórico y teorías explícitas distintas relacionadas que la hacen más fuerte), vale la pena diferenciar si el origen es uno u otro.

Un tercer candidato es el de una *teoría tácita*, entendiendo ésta como una teoría inconciente que subyace al desempeño en alguna tarea específica dentro de un campo específico. La postulación de teorías tácitas bien puede rastrearse en la propuesta de la gramática universal de Chomsky (1959): “Chomsky argumentó que, codificadas en los cerebros de los seres humanos normales, hay representaciones de las gramáticas de los lenguajes naturales que ellos hablan. Aquellas representaciones apuntalan nuestra capacidad cotidiana de entender y producir oraciones gramaticales” (Ravenscroft 2009, p. 135)³. La gramática universal chomskiana funciona como una teoría tácita, los hablantes competentes de un lenguaje si bien la ocupan en sus enunciaciones, no tienen acceso a ella, así que no es posible vía introspección conocer las reglas que subyacen al desempeño lingüístico. Esta estrategia ha sido usada, con éxito diferencial, para explicar el desempeño en otros campos, como el moral (Hauser 2006) o la comprensión básica de algunos fenómenos cognitivos tal como la psicología popular (Greenwood 1991) o la biología popular (Medin & Atran 1999).

Cabe señalar dos aspectos importantes. Por un lado, algunas de estas gramáticas o teorías tácitas no tienen un valor de verdad, tal es el caso de la lingüística y, tal vez, la moral, mientras que otras, como la psicología y la biología popular, bien pueden estar acertadas o erradas respecto al fenómeno que versan. Como se desarrollará posteriormente (ver §1.4.1.), la aparición de teorías tácitas puede deberse a una estrategia rápida y exitosa de resolver problemas que no necesariamente está comprometida con la verdad. Por otro lado, comprometerse con la existencia de una teoría tácita no implica comprometerse con todas. Por ejemplo, se puede defender la existencia de una teoría tácita acerca de los seres vivos sin defender la existencia de una gramática moral. Las discusiones acerca de la existencia de cada una, su utilidad explicativa, y los compromisos que implican, serán independientes.

³ “Chomsky argued that encoded in the brains of normal human beings are representations of the grammars of the natural languages which they speak. Those representations underpin our everyday capacity to understand and produce grammatical sentences” (Ravenscroft 2009, p. 135).

El siguiente cuadro ilustra la historia de las fuentes de las intuiciones que he presentado en este apartado. El esquema no busca modelar un proceso cognitivo de surgimiento de intuiciones, más bien pretende establecer las relaciones de éstas con las posibles fuentes.

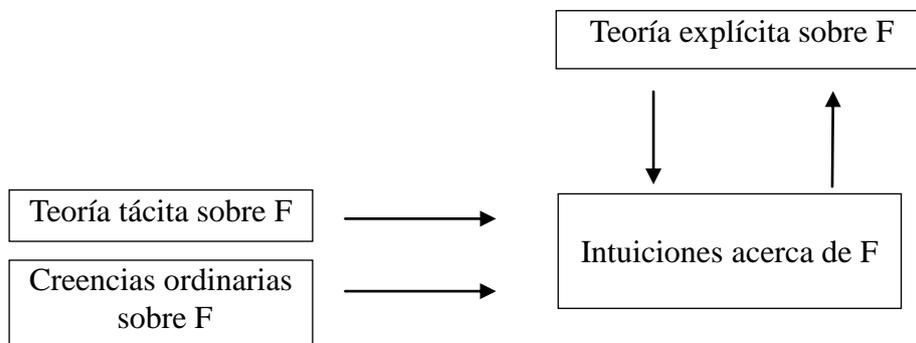


Fig. 1. Modificado a partir de Cummins (1998, p. 122)

Enfrentado a un escenario relacionado con F, un sujeto podría tener intuiciones derivadas tanto por una teoría popular tácita, que haga parte de su andamiaje psicológico, pero también a partir de una creencia fruto de su historia personal. Sin embargo, si dicho sujeto posee una teoría explícita sobre F, lo más probable será que ésta afecte sus intuiciones y reemplace el insumo de una teoría tácita o una creencia ordinaria, dado el valor epistémico más alto que, se espera, el sujeto le atribuye. En particular, esto será verdad en caso de sujetos interesados por las teorías explícitas, tales como académicos, científicos o filósofos.

De esta manera, las intuiciones populares, serán aquellas que sean producidas por teorías tácitas. El carácter contextual de las creencias ordinarias y, más aún, de las teorías explícitas hacen que las intuiciones que produzcan no brinden confianza acerca de su estabilidad entre contextos diferentes.

§1.2. Cuatro tesis de la Filosofía Experimental

Teniendo presente la anterior caracterización sobre las intuiciones, presentaré en este apartado las tesis metafilosóficas, filosóficas y metodológicas que proveen el trasfondo a partir del cual los filósofos experimentales desarrollan su trabajo. Propongo cuatro tesis básicas compartidas por los filósofos experimentales y que son definitivas para enmarcar el trabajo que llevan a cabo.

Las tesis son:

- I. Existen problemas filosóficos tales que en su resolución deben tomarse en cuenta las intuiciones populares.
- II. Cualquier apelación a intuiciones populares tiene un fuerte compromiso empírico.
- III. El método filosófico usual para acceder a las intuiciones populares no está a la altura del compromiso empírico.
- IV. Los tests basados en encuestas son adecuados para la tarea de sistematizar las intuiciones populares.

A continuación presentaré cada una de estas tesis y en su enunciación iré introduciendo y aclarando algunos de los términos usados. En la sección §1.4. revisaré las controversias que han surgido alrededor de éstas.

§1.2.1. Tesis I: Existen problemas filosóficos tales que en su resolución deben tomarse en cuenta las intuiciones populares

Ésta es una tesis metafilosófica esencial al programa de la Filosofía Experimental. En la literatura se muestra cierta preocupación por evitar un *parroquianismo* (Nichols 2004a) que haga que los filósofos sólo discutan acerca de conceptos técnicos usados por otros filósofos. Esta preocupación se debe a que si la Filosofía sólo tratara sobre el contenido de aquellos conceptos técnicos, entonces, “estaría divorciada de las preocupaciones iniciales que llevaron a la investigación filosófica del concepto y sus veredictos tendrían poca influencia en esas preocupaciones iniciales. Como tal, una porción grande y central de la práctica filosófica debe estar versada sobre conceptos ordinarios” (Alexander & Weinberg 2007, 57)⁴. Por supuesto, no se pretende desestimar las discusiones especulativas acerca de conceptos técnicos que se hallan muy alejados de los conceptos populares. Es más interesante establecer una cierta diferencia entre problemas filosóficos, que permita acotar el abanico de problemas en los que el enfoque de la Filosofía Experimental puede ser útil.

Por un lado, es útil hacer la distinción entre problemas filosóficos que tienen origen en fenómenos ya conceptualizados cotidianamente (o fácilmente conceptualizables desde el sentido común) y aquellos que tienen origen en el seno de las discusiones técnicas o a partir de hallazgos científicos.

⁴ “[...] it would be divorced from the concerns that led us to philosophical investigation of the concept in the first place and its verdicts would have little bearing on those initial concerns. As such, large and central swaths of philosophical practice must be concerned with the ordinary concepts” (Alexander & Weinberg 2007, 57)

Para los primeros, las intuiciones juegan un papel muy relevante dado que la sistematización de aquellas permite delimitar el *explanandum* de aquello que se quiere explicar, por otra parte tienen un papel evidencial en la construcción de argumentos a favor y en contra de teorías al respecto. Para los segundos, el papel de las intuiciones es nulo dado que, o la discusión no tiene relación con algún concepto ordinario, y por tanto la apelación a ellos no tiene valor, o la discusión se enriquece con evidencia de las ciencias naturales que tendrían un papel predominante frente a las intuiciones. El peso que se le da a las intuiciones variará respecto al problema filosófico en cuestión. Desde aquel, por ejemplo, que versa sobre la naturaleza de la “materia” en el que la evidencia empírica juega un papel fundamental y no importa, incluso es de esperarse, si una teoría acerca de ella es calificada como contraintuitiva, hasta, por ejemplo, el problema del “conocimiento”, donde una teoría que no esté de acuerdo con las intuiciones será desechada. Por supuesto, entre estos dos extremos se incluyen algunos casos intermedios problemáticos.

Pero tomar en cuenta las intuiciones populares con respecto a un fenómeno no es afirmar que dichas intuiciones populares constituyan una respuesta a las preguntas que desde una disciplina se formulen sobre dicho fenómeno. Las intuiciones no deben considerarse fuente de verdades (Cummins 1998; Wisniewsky 1998), esto sólo sería producto de un ingenuo innatismo divino de corte cartesiano⁵ o del más exagerado evolucionismo panglossiano⁶. Las intuiciones jugarán papeles distintos en el proceso de resolución de los problemas relacionados con dichos fenómenos, algunas veces servirán como materia prima a partir de la cual plantear hipótesis y en otras como evidencia (cuyo valor, impacto y grado dependerá del problema en cuestión) para aceptar o rechazar tesis.

§1.2.2. Tesis II: Cualquier apelación a intuiciones populares tiene un fuerte compromiso empírico

Dada la definición de “intuición popular” que he ofrecido:

el juicio espontáneo, fruto de un proceso no accesible de manera conciente, compartido por la gran mayoría de los sujetos, normales y competentes del lenguaje, que no posean (o no hayan tenido contacto con) una teoría propia de la disciplina que estudia el fenómeno sobre el que versa el juicio. (ver §1.1.)

⁵ Aquel que afirme la veracidad de las intuiciones en tanto que ideas innatas proveídas por una divinidad que garantiza su verdad.

⁶ Aquel que afirme la veracidad de las intuiciones en tanto que ideas innatas parte del precableado cerebral fruto de la evolución y por tanto veraces dado que permitieron la supervivencia de los hombres en los que se afianzaron como rasgos.

Apelar a intuiciones populares parece implicar un compromiso empírico. Si un filósofo afirma que “es intuitivo que x ” o usa expresiones similares como “dado que es común pensar que y ” está haciendo una afirmación acerca de un hecho del mundo: que dicha creencia sea compartida, de manera explícita o tácita, por la gran mayoría de sujetos normales competentes del lenguaje.

§1.2.3. Tesis III: El método filosófico usual para acceder a las intuiciones populares no está a la altura del compromiso empírico

Respecto a cómo un filósofo consulta las intuiciones para su argumentación es muy ilustrativa la afirmación de Jackson (1998, p. 37): “con frecuencia sabemos que nuestro propio caso es típico y así podemos generalizarlo a los demás”. De hecho, la idea de que establecer las intuiciones puede hacerse con el método de considerar la propia intuición y luego generalizarla es una idea de vieja data en la Filosofía⁷.

Sin embargo, este método puede no estar a la altura del compromiso empírico que implica apelar a intuiciones populares. Al ser la intuición fruto de un proceso al que no se accede de manera conciente, la o las fuentes que generan una intuición particular son inaccesibles a quien intuye algo. Sopesar clara y distintamente qué tanto de creencia ordinaria, de teoría implícita o de teoría explícita tiene la propia intuición es una tarea muy complicada. Si las intuiciones de las que el filósofo se vale son producto de creencias ordinarias, sólo serán una muestra de la historia particular del filósofo y del acervo de creencias compartidas por su comunidad, así no tendrá la fuerza para apoyar un argumento. De igual forma, si la intuición en cuestión es fruto de una teoría explícita, la fuerza de la argumentación recaerá en la teoría que genera la intuición y si bien ésta puede ser correcta sigue siendo una intuición eminentemente contextual que, en principio, no puede tener pretensiones de generalidad. En este caso será mejor que el filósofo apele a la fuerza y valor de la teoría que genera dicha intuición y no a la intuición misma.

Así, el método descrito por Jackson (1998), aceptado y usado por la mayoría de filósofos analíticos contemporáneos, no ofrece una total confianza respecto al origen de la intuición del filósofo. Hay buenos motivos para sospechar que la intuición del filósofo, que luego es generalizada, puede ser fruto de sus teorías explícitas o de sus creencias ordinarias, relativas a su historia personal como

⁷ Un solo ejemplo. El siguiente es un extracto de la introducción del *Leviatán* de Hobbes: “[...] por la semejanza de los pensamientos y de las pasiones de un hombre con los pensamientos y pasiones de otro, quien se mire a sí mismo y considere lo que hace cuando *piensa, opina, razona, espera, teme, etc.*, y por qué razones, podrá leer y saber, por consiguiente cuáles son los pensamientos y pasiones de los demás hombres en ocasiones parecidas”. (Hobbes 2008, p. 4)

vimos anteriormente. Así, se abre el espacio a un método adicional que ofrezca mayor confianza respecto al compromiso empírico en la apelación a intuiciones.

§1.2.4. Tesis IV: Los tests basados en encuestas son adecuados para la tarea de establecer intuiciones populares

Esta tesis metodológica es la que hace experimental a la Filosofía Experimental. Los filósofos experimentales ejecutan investigaciones empíricas acerca de las intuiciones populares de los sujetos, siendo éste el rasgo característico del movimiento (Knobe & Nichols 2008; Nahmias et. al. 2005; Alexander & Weinberg 2007). Estas investigaciones se hacen usando los métodos típicos de la Psicología Experimental: experimentación sistemática y análisis estadístico inferencial.

La metodología de más uso, como muestran los artículos en Filosofía Experimental, consiste en aplicar pruebas basadas en encuestas a una muestra relevante de sujetos. El diseño de las pruebas busca individuar y establecer las intuiciones populares de los sujetos respecto a ciertos fenómenos una vez se encuentre una tendencia estadísticamente significativa en las respuestas ofrecidas por ellos.

Hay varias consideraciones respecto a privilegiar este tipo de metodología que sólo valora la evidencia lingüística. Otras pruebas en Psicología Experimental, donde se toma en cuenta evidencia neurobiológica o evidencia conductual, no serán de mucha utilidad para individuar y especificar las intuiciones populares de los sujetos. Si bien la evidencia neurobiológica en algunas ocasiones puede resultar útil para ciertas investigaciones en Filosofía Experimental⁸, en el mejor de los casos sólo ayudará a esclarecer los procesos neurofisiológicos que subyacen a la generación de las intuiciones; esto puede ser importante para una tarea descriptiva (ver §1.4.2.), pero no servirá para especificar de manera sustantiva las intuiciones populares. Dado que éstas, en tanto que generadas por teorías tácitas, se encuentran en un nivel cognitivo y no tendrán un tipo de correlato neuronal con el que se puedan relacionar uno a uno.

⁸ Un caso ilustrativo es el de la investigación de Koenigs et. al. (2007), donde aplican un cuestionario con dilemas morales personales (aquellos donde hay contacto directo con víctimas) y dilemas impersonales (aquellos donde no hay contacto directo con víctimas) a pacientes con daños en la corteza frontal, para poner a prueba ciertas hipótesis a favor de un modelo dual de la construcción del juicio moral (donde juegan un rol importante tanto los estados cognitivo/racionales y los estados de emociones). Los resultados mostraron que aquellos con un daño en la corteza frontal tenderán a tener respuestas utilitaristas en los dilemas personales, mientras que los sujetos normales tienen respuestas provocadas por las emociones negativas que se producen en dicha área del cerebro.

Las pruebas en Psicología Experimental que se ocupan de recolectar evidencia conductual tampoco parecen resultar adecuadas para el fin de caracterizar las intuiciones populares. Podría pensarse que habrá algunas intuiciones populares relacionadas con las acciones de los sujetos y que revisando estas acciones se podrían ubicar dichas intuiciones. Sobre este punto deben hacerse dos consideraciones: i) el abanico de fenómenos sobre los que puedan evaluarse las acciones buscando encontrar las intuiciones populares que están detrás de ellas es muy limitado; ii) en los experimentos en Psicología donde los sujetos deben enfrentarse a una tarea y actuar de cierto modo, los factores irrelevantes que pueden generar las respuestas conductuales pueden ser más, y más difíciles de evaluar, que en el caso de la discriminación de factores irrelevantes que provocan intuiciones populares en la resolución de encuestas.

Recuérdese el célebre experimento mental, donde se quieren probar las intuiciones morales. En él, se plantea un escenario en el que un tren va directo a arrollar a cinco personas. El sujeto tiene cerca una palanca que de accionarla cambiaría el curso del tren hacia una dirección donde sólo arrollará a una persona (Foot 1967). Si alguien tuviese intuiciones utilitaristas se espera que opte por accionar la palanca dado que desde el utilitarismo una muerte será preferible a cinco. Si este experimento se hiciera en vivo, con tren, palanca y víctimas, muy seguramente además de la intuición utilitarista entrarían en juego muchos otros factores: la sangre fría del sujeto, el coraje, el nivel de respuesta en situaciones de conflicto, entre otros que no le interesan evaluar al filósofo experimental en este caso particular. Así, el estudio de las intuiciones que permite la encuesta aísla muchos factores pragmáticos que entrarían en juego en estudios conductuales.

Ahora bien, siendo que la metodología de cuestionarios de preguntas y respuestas luce adecuada, se abre una disyuntiva acerca del tipo de preguntas: ¿preguntas abiertas o preguntas cerradas?, la respuesta a esa pregunta dependerá de qué tanto se quiera privilegiar un método más o menos fácil de cuantificar. Hay cierto consenso general respecto a que las preguntas más fáciles de cuantificar son aquellas con estructura completa, donde el sujeto está obligado a responder dentro de un número reducido de respuestas posibles. En las preguntas abiertas, aquellas semi-estructuradas, el sujeto tendrá la responsabilidad de hacer un trabajo interpretativo, donde además de ocuparse de qué responder, deberá ocuparse de qué se espera que responda. Esta tarea interpretativa amplía la entrada de ciertos factores contextuales sobre la manera en que el sujeto responderá la pregunta. Dado que es deseable que el sujeto responda sin atender a factores contextuales y relativos, se preferirán aquellas preguntas donde toda la información disponible al sujeto esté en el escenario descrito y donde las posibles respuestas estén especificadas. Otra estrategia experimental usada para

identificar y disminuir los factores contextuales es ejecutar la prueba en diversos contextos para establecer el peso de ellos y así lograr aislarlos.

§1.3. Algunos ejemplos

Esta elección metodológica bebe de varias fuentes de la investigación psicológica (Nichols 2004a), así la investigación empírica de los filósofos experimentales puede enmarcarse, aunque no exclusivamente, dentro de la Psicología Social, la Psicología Transcultural o la Psicología del Desarrollo. A continuación presentaré algunas experiencias en Filosofía Experimental que se podrían ubicar dentro de las áreas antes mencionadas. No discutiré el éxito de estos trabajos, los presentaré para ejemplificar la tesis metodológica de los filósofos experimentales. Los tres difieren tanto en el acercamiento utilizado como en los temas: teoría de la acción, referencia de nombres propios y libre albedrío.

§1.3.1. Knobe, J. (2004) - Psicología Social

Dentro de las discusiones generales en relación con la acción intencional, se encuentra la pregunta acerca de si los efectos secundarios de las acciones pueden contarse como intencionales o no. Joshua Knobe (2004) ingresa a la discusión ofreciendo evidencia empírica de las intuiciones populares acerca de acciones intencionales en casos complejos como el que revisaré a continuación.

El presidente de una compañía ha decidido implementar un programa que hará que la compañía gane mucho dinero, además cree que la implementación tendrá un efecto secundario x . Al presidente no le importa el efecto secundario e implementa el nuevo programa esperando mejorar las ganancias de la compañía. Una vez implementado, todo ocurre de acuerdo a lo previsto; la compañía mejora sus ganancias y el efecto secundario x , sucede. ¿Puede decirse que el presidente produjo este efecto secundario intencionalmente?

Knobe, luego de hacer una revisión de la literatura donde se sostienen diferentes posiciones respecto a la pregunta anterior, ofrece evidencia proveniente de pruebas en Psicología Social basadas en encuestas donde encuentra un hallazgo interesante: la intuición popular califica como intencional un efecto secundario si este es considerado inmoral, si lo contrario es el caso, es decir, si el efecto secundario es considerado moral, la intuición popular no lo califica como intencional.

El experimento principal se realizó sobre una población de 78 sujetos que se encontraban en un parque público de Manhattan, a cada uno de los sujetos se le asignó, de manera aleatoria, la “condición de daño” o la “condición de ayuda”. Los sujetos en la “condición de daño” leyeron la siguiente viñeta:

El vicepresidente de una compañía buscó al presidente y le dijo: “Estamos pensando en comenzar un nuevo programa. El incrementará nuestras ganancias, pero también dañará el medio ambiente.

El presidente respondió: “No me importa en absoluto el daño al ambiente. Sólo quiero aumentar las ganancias tanto como pueda. ¡Comencemos el nuevo programa!

Ellos comenzaron el programa. Por supuesto, el medio ambiente recibió daños. (Knobe 2003, p. 190)⁹

A continuación se le solicitó a los sujetos que evaluaran la responsabilidad del presidente [*chairman*] (en una escala de 0 a 6) y que respondieran a la pregunta ¿Usted cree que el presidente dañó intencionalmente el ambiente?

Los sujetos en la “condición de ayuda” recibieron una viñeta idéntica, sólo que la palabra “daño” [*harm*] fue reemplazada por “ayuda” [*help*]. Así:

El vicepresidente de una compañía buscó al presidente y le dijo: “Estamos pensando en comenzar un nuevo programa. El incrementará nuestras ganancias, pero también ayudará el medio ambiente.

El presidente respondió: “No me importa en absoluto ayudar al medio ambiente. Sólo quiero aumentar las ganancias tanto como pueda. ¡Comencemos el nuevo programa!

⁹ “The vice-president of a company went to the chairman of the board and said, ‘We are thinking of starting a new program. It will help us increase profits, but it will also harm the environment.’

The chairman of the board answered, ‘I don’t care at all about harming the environment. I just want to make as much profit as I can. Let’s start the new program.’

They started the new program. Sure enough, the environment was harmed” (Knobe 2003, p. 190).

Ellos comenzaron el programa. Por supuesto, el medio ambiente recibió ayuda. (Knobe 2003, p. 190)¹⁰

Los resultados mostraron que las dos condiciones provocaron patrones de respuestas muy diferentes, bajo la “condición de daño” el 82% de los sujetos respondió que el agente produjo el efecto de manera intencional, mientras que bajo la “condición de ayuda” el 77% de los sujetos respondió que el efecto se produjo de manera no intencional. Luego de mostrar cómo estas diferencias son estadísticamente significativas, Knobe (2003) establece una correlación entre las respuestas ofrecidas por los sujetos a la pregunta por la intencionalidad del efecto secundario y la evaluación de la responsabilidad del presidente respecto al efecto secundario. En la evaluación de la responsabilidad también se presenta una asimetría que sugiere una teoría que puede explicar la asimetría en la atribución de intencionalidad en el caso de los efectos secundarios: “ellos parecen considerablemente más dispuestos a decir que un efecto secundario fue provocado intencionalmente cuando el resultado del efecto secundario es malo que cuando es bueno” (Knobe 2003, p. 193).¹¹

Aunque la muestra de los sujetos del experimento de Knobe (2003) es pequeña y además todos comparten un trasfondo cultural muy similar, los resultados han sido replicados (McCann 2005, Machery 2008), así se ha ratificado que sí hay un efecto interesante y un rompecabezas por resolver. La evidencia presentada por Knobe no soluciona definitivamente el problema de la intencionalidad de los efectos secundarios, pero ofrece una restricción empírica que supone un avance interesante: es un error esperar una respuesta general a la pregunta acerca de cómo las personas califican de intencional o no un efecto secundario. Por otra parte, muestra un asunto interesante: el uso de un concepto que se esperaba fuera puramente descriptivo, como lo es “intencional”, está atravesado por una consideración normativa de lo “bueno” y lo “malo” (Uttich & Lombrozo en prensa) .

Knobe (2003) logró ingresar el componente de la “calificación moral” al análisis de la acción intencional y, tal como señala Nichols (2004a), aunque las conclusiones y el diseño experimental de Knobe (2003) puede ser criticado, resultará más fácil evaluar dichas críticas que dirimir entre intuiciones contrarias presentadas por filósofos con un enfoque tradicional. En otras palabras, si un

¹⁰ “The vice-president of a company went to the chairman of the board and said, ‘We are thinking of starting a new program. It will help us increase profits, but it will also help the environment.’

The chairman of the board answered, ‘I don’t care at all about helping the environment. I just want to make as much profit as I can. Let’s start the new program.’

They started the new program. Sure enough, the environment was helped” (Knobe 2003, p. 190).

¹¹ “they seem considerably more willing to say that a side-effect was brought about intentionally when they regard that side-effect as bad than when they regard it as good” (Knobe 2003, p. 193).

filósofo pretendía argumentar la intencionalidad de un efecto colateral a partir de sus intuiciones al respecto mientras que otro argumentara en contra, basado en sus propias intuiciones, la discusión acerca de qué intuiciones de cuál filósofo serán más valiosas podrá zanjarse señalando cuáles son intuiciones populares y cuáles no.

§1.3.2. Machery, E., Mallon, R., Nichols, S., y Stich, S. (2004) – Psicología Transcultural

Es celebre el experimento mental de Kripke (1980) en contra de la teoría descriptivista y a favor de una teoría histórico-causal de la referencia:

Supóngase que Gödel no fue de hecho el autor del teorema. Un hombre llamado Schdmit realmente hizo el trabajo en cuestión. Su amigo Gödel tomó el manuscrito que posteriormente se le atribuyó a Gödel. Bajo la teoría en cuestión cuando un hombre ordinario use el nombre “Gödel”, él realmente refiere a Schdmit, porque Schdmit es la única persona que satisface la descripción “el hombre que descubrió la incompletitud de la aritmética”.

La intuición a la que apela Kripke es que a pesar de que Schdmit es la única persona que satisface la descripción de “Gödel”, ese nombre no refiere a Schdmit. Así, una teoría descriptivista no ofrece un resultado deseable dado que resulta contraintuitivo afirmar que “Gödel” refiere a Schdmit.

Machery et. al. (2004) pusieron a prueba la intuición a la que apela Kripke en su argumentación a favor de una teoría histórico-causal, por medio de estudios en Psicología Trans-Cultural, evaluando si efectivamente esta intuición es sostenida por sujetos Occidentales y por sujetos de Asia Oriental.

A los sujetos se les presentó la siguiente versión del caso de Kripke (1980) y se les formuló una pregunta:

Supón que John aprendió en la escuela que Gödel es el hombre que probó un importante teorema matemático, llamado la incompletitud de la matemática. John es muy bueno en matemáticas y puede dar una instrucción precisa del teorema de la incompletitud, además de atribuirle su descubrimiento a Gödel Pero esta es la única cosa que el ha escuchado acerca de Gödel. Ahora supón que Gödel no fue el autor de este teorema. Un hombre llamado “Schdmit”, cuyo cuerpo fue encontrado en Viena bajo misteriosas circunstancias muchos años atrás, en realidad fue quien hizo el trabajo en cuestión. Su amigo Gödel de algún modo se apoderó del manuscrito y reclamó el crédito por el trabajo que a partir de entonces fue

atribuido a Gödel. Así, él ha sido conocido como el hombre que probó la incompletitud de la aritmética. La mayoría de las personas que han escuchado el nombre “Gödel” son como John; la afirmación de que Gödel descubrió el teorema de la incompletitud es la única cosa que ellos han escuchado acerca de Gödel.

Cuando John usa el nombre “Gödel”, él está hablando acerca de:

(A) la persona que realmente descubrió la incompletitud de la aritmética

o

(B) la persona que se apoderó del manuscrito y reclamó el crédito por el trabajo

(Machery et. al. 2004, p. B6)¹²

Los resultados ofrecidos por Machery et. al. (2004) mostraron que mientras los Occidentales tendieron a responder (B) conforme a las teorías histórico-causales de la referencia, y a la intuición apelada por Kripke, los sujetos de Asia Oriental tendieron a responder (A) conforme a las teorías descriptivistas de la referencia. Esta evidencia, según Machery y sus colegas, motiva un rechazo al argumento de Kripke (1980) en contra de las teorías descriptivistas de la referencia, dado que descansa sólo sobre intuiciones Occidentales.

El trabajo de Machery et. al. (2004) ha sido polemizado en varias ocasiones. Algunos han señalado que su experimento sólo muestra intuiciones metalingüísticas y no del uso de nombres, no pudiendo así establecer un punto fuerte en contra del argumento de Kripke (Martí 2009), mientras otros han hecho experimentos que arrojan evidencia en contra (Lam en preparación), donde hablantes de cantonés sostienen intuiciones populares histórico-causales respecto a la referencia de los nombres, desafiando los resultados de Machery et. al. (2004). Aún sin conocer los resultados finales del debate, la investigación de Machery et. al. (2004) sugiere una conclusión interesante: mostrar la

¹² “Suppose that John has learned in college that Gödel is the man who proved an important mathematical theorem, called the incompleteness of arithmetic. John is quite good at mathematics and he can give an accurate statement of the incompleteness theorem, which he attributes to Gödel as the discoverer. But this is the only thing that he has heard about Gödel. Now suppose that Gödel was not the author of this theorem. A man called “Schmidt”, whose body was found in Vienna under mysterious circumstances many years ago, actually did the work in question. His friend Gödel somehow got hold of the manuscript and claimed credit for the work, which was thereafter attributed to Gödel. Thus, he has been known as the man who proved the incompleteness of arithmetic. Most people who have heard the name “Gödel” are like John; the claim that Gödel discovered the incompleteness theorem is the only thing they have ever heard about Gödel.

When John uses the name “Gödel”, is he talking about:

(A) the person who really discovered the incompleteness of arithmetic?

or

(B) the person who got hold of the manuscript and claimed credit for the work?” (Machery et. al. 2004, p. B6)

variación transcultural de las intuiciones populares socava la autoridad de los filósofos al teorizar *a priori* acerca de conceptos como “referencia” ya que ellos, como cualquier otro, no tienen acceso a los aspectos culturalmente locales de sus intuiciones (Nichols 2004a, Knobe & Nichols 2008).

§1.3.3. Nichols, S. (2004b) – Psicología del Desarrollo

Una visión tradicional respecto al libre albedrío, es la teoría de agente-*causa*, según la cual un agente realizó una acción libremente si el agente pudo no haberla realizado. Nichols (2004b) ofrece evidencia empírica del soporte de esta teoría en intuiciones populares compartidas por niños y adultos.

En los experimentos ejecutados por Nichols (2004b) los niños, de una edad media de 4 años, fueron separados de acuerdo a dos condiciones distintas. Aquellos en la “condición agente” fueron testigos de un agente humano exhibiendo comportamiento motor y aquellos en la “condición objeto” fueron testigos del movimiento de un objeto. A los niños se les preguntó si el agente/objeto tenía que comportarse de esa manera o si pudo haber hecho alguna otra cosa. Los resultados mostraron que los niños consideran en el caso del agente humano que dicho agente pudo haberse comportado de manera diferente, mientras que en el caso del objeto consideran que no pudo haberse comportado de manera diferente.

En otros experimentos, Nichols (2004b) presentó a grupos de adultos y niños escenarios de sucesos físicos, e.g. una olla de agua que llega al punto de ebullición, y de sucesos que implican una elección moral, e.g. una niña robando en una dulcería. A los participantes se les preguntó: si todo en el mundo fuera idéntico hasta el momento en el que el evento ocurrió, ¿el suceso *tenía* que ocurrir? El resultado mostró que ambos, adultos y niños, tienden a considerar que los sucesos físicos tenían que ocurrir, mientras que los sucesos que implican una elección moral no *tenían* que ocurrir.

Según Nichols (2004a, 2004b), esto muestra evidencia de que las intuiciones populares favorecen una teoría de agente-*causa* respecto al problema del libre albedrío. Una teoría diferente que proponga otro acercamiento podría ser calificada de *contraintuitiva* y tendría que ofrecer respuestas convincentes acerca de por qué abandonar tales intuiciones.

§1.4. Alcances de la Filosofía Experimental

En el apartado anterior presenté las tesis centrales de la Filosofía Experimental y di un repaso breve a algunos trabajos, con el fin de ejemplificar su *modus operandi*. En este apartado discutiré algunos de los alcances y límites de este enfoque con respecto a problemas filosóficos clásicos. Señalarlos permitirá responder a la pregunta acerca de qué podemos esperar, y qué no, de los resultados de este enfoque.

§1.4.1. El abanico de problemas es limitado

Tal como sugiere la Tesis I (ver §1.2.1.), los problemas filosóficos a los que el enfoque de la Filosofía Experimental puede aplicarse esperando resultados interesantes se limita a aquellos que versan sobre fenómenos ya conceptualizados, o fácilmente conceptualizables, en el sentido común. Un criterio que considero útil para establecer esta distinción es el siguiente: un debate versa sobre fenómenos de este tipo, toda vez que una teoría presentada sobre el fenómeno puede ser descalificada por resultar contraintuitiva. Este criterio es bastante débil y podría señalarse como circular, pero, puede tener un valor heurístico para identificar el tipo de problemas a los que estoy haciendo referencia.

No se espera que las grandes preguntas filosóficas tengan respuesta en las intuiciones populares, éste sería un objetivo ingenuo. Las teorías filosóficas que pueden tener cierto impacto por la investigación a partir de este enfoque son aquellas que se basan en i) experimentos mentales, ii) referencias al sentido común o a intuiciones pre-filosóficas, iii) análisis conceptual basado, en parte, en el uso ordinario de los términos (Nadelhoffer & Nahmias 2007). Las discusiones filosóficas en donde la Filosofía Experimental no tiene mucho que aportar, serán aquellas que versen sobre i) conceptos técnicos que no tengan ni origen ni equivalente en el uso cotidiano o ii) clases naturales sobre los que se puedan hacer teorías explícitas en ciencias naturales.

En las teorías basadas en experimentos mentales el papel de las intuiciones es central. Este tipo de argumentos¹³ reciben su nombre dado que comparten ciertos aspectos con los experimentos llevados a cabo en la ciencia natural. Por un lado, hacen un cambio controlado de variables, establecen cuáles variables dependen de otras por medio de hipótesis y las someten a prueba

¹³ Partiré de la aproximación según la cual los experimentos mentales son argumentos o que al menos pueden reconstruirse como tales. Esta posición deja más o menos abierta la cuestión del papel de la imaginación en los experimentos mentales.

(Brendel 2003. p. 10). Como es claro, la idea misma de “poner a prueba” implica que las intuiciones, como los hechos observados en el caso de los experimentos científicos, son evidencia que contrasta las teorías. Así, debe haber claridad si las intuiciones populares a las que apelan estos argumentos son en realidad tales.

En el caso de las teorías donde se hace referencia al sentido común o a intuiciones pre-filosóficas, es aún más clara la manera en el que la Filosofía Experimental puede hacer aportes interesantes. Casos claros son aquellos en los que las intuiciones hacen parte de las premisas de un argumento filosófico sofisticado. Recuérdese el argumento de Nagel (1974) a favor del carácter esencialmente subjetivo de la experiencia conciente. Parte de la argumentación descansa en una premisa explícita que aspira a ser intuitiva: “Supongo que todos creemos que los murciélagos tienen experiencias” (Nagel 1974). El aparato argumentativo de Nagel en contra del acceso intersubjetivo a las experiencias de criaturas de un tipo distinto al nuestro no podría funcionar si aquello que supone fuera falso, es decir, si no fuera una intuición popular que los murciélagos tienen experiencias. Este punto generalmente no ha sido considerado como indispensable para el argumento de Nagel, y no es mi intención motivar esa lectura, sólo me interesa resaltar de qué manera apelar a una intuición como punto de partida para una argumentación filosófica es otro de sus posibles usos, y por tanto allí hay un espacio donde la investigación empírica sobre las intuiciones populares puede ser relevante.

Lo mismo ocurre para tareas de análisis conceptual donde se parte del uso ordinario de conceptos. Allí donde un filósofo señale “el uso ordinario de un concepto x es...” se está asumiendo el compromiso empírico señalado en la Tesis II (ver §1.2.2.), lo cual supone que el trabajo empírico sobre los usos ordinarios de los conceptos puede ser importante. Por otro lado, si como señala Sorensen (1998) “Los filósofos analíticos hacen un uso frecuente de experimentos mentales porque es una prueba natural para las prácticas clarificadoras que constituyen al análisis conceptual: definición, delegación de preguntas, trazar distinciones, construir condiciones de adecuación, explicitar vínculos, adelantar pruebas de posibilidad, mapear patrones de inferencia” (Sorensen 1998, p. 15)¹⁴, entonces, allí donde se usen experimentos mentales en las tareas de análisis conceptual también la Filosofía Experimental tendrá relevancia.

¹⁴ “Analytic philosophers make heavy use of thought experiment because it is the natural test for the clarificatory practices constituting conceptual analysis: definition, question delegation, drawing distinctions, crafting adequacy conditions, teasing out entailments, advancing possibility proofs, mapping inference patterns” (Sorensen 1998, p. 15).

§1.4.2. Tareas disponibles desde la Filosofía Experimental

El alcance de un estudio en Filosofía Experimental puede establecerse a partir de tres tareas básicas que se encuentran disponibles. Los datos de las intuiciones populares, pueden usarse para: i) soportar teorías filosóficas, ii) mostrar que dichas intuiciones no pueden servir de apoyo para las teorías filosóficas que las han usado, y iii) entender los mecanismos psicológicos que generan tales intuiciones (Nadelhoffer & Nahmias 2007).

Nadelhoffer & Nahmias (2007) han establecido una taxonomía de los trabajos en Filosofía Experimental de acuerdo al uso que se le da a los resultados de las investigaciones acerca de intuiciones populares; encuentro útil la clasificación, toda vez que permite identificar más claramente los objetivos y alcances esperados:

Análisis Experimental (o programa positivo). Uso del estudio de intuiciones para apoyar teorías filosóficas. No choca con la tesis según la cual las intuiciones suponen evidencia para una teoría filosófica, en cambio difiere acerca de la metodología para obtener dichas intuiciones. El caso presentado de investigación en Psicología del Desarrollo sobre intuiciones populares acerca de libre albedrío (Nichols 2004a, 2004b) cae en esta categoría.

Restriccionismo Experimental (o programa negativo). Uso del estudio de intuiciones como una batalla en contra del uso (y abuso) de intuiciones en la Filosofía. Sostiene la tesis de que si nuestras intuiciones acerca de un tema particular varían trans-culturalmente o socioeconómicamente y no tenemos cimientos independientes para privilegiar nuestras intuiciones por encima de aquellas sostenidas por otros, entonces esas intuiciones son insuficientes para construir una teoría filosófica. El trabajo de Machery et. al. (2004) sobre referencia de nombres propios se ubicaría en este rubro.

Descriptivismo Experimental. Uso del estudio de intuiciones para entender los mecanismos psicológicos tras la generación de esas intuiciones, bien podría también decirse que se ocupa de la sistematización de las teorías tácitas subyacentes. Típicamente se varían los escenarios para identificar aquellos aspectos relevantes que influyen la producción de intuiciones. Esta tarea descriptiva puede arrojar luces acerca de la arquitectura cognitiva humana (Nichols 2004a).

Es importante hacer algunas notas respecto a los dos primeros programas. Hacer el contraste entre el programa positivo y el programa negativo será importante para dimensionar el impacto de las críticas que se presentarán en el siguiente apartado. Los dos programas están de acuerdo en todas las cuatro tesis generales. Sin embargo, difieren en su tratamiento de las intuiciones una vez que son estudiadas. Los trabajos realizados en el programa negativo (e.g. Machery et. al. 2004) buscan evidencia a favor de la inestabilidad de las intuiciones relevantes a problemas filosóficos, mostrando que tal variación se debe a factores contextuales no relevantes filosóficamente. El programa negativo procura ser un desafío de la llamada “filosofía de sillón”, aquella que a partir del análisis conceptual y las propias intuiciones del filósofo construyen teorías. En tal programa hay un abanico de posiciones, desde los que hacen una denuncia respecto a la universalidad de algunas intuiciones sobre algún experimento mental, hasta los que consideran tener evidencia suficiente para ser totalmente escépticos de la utilidad de las intuiciones en alguna empresa filosófica. Por el contrario, los trabajos realizados en el programa positivo suponen la estabilidad de las intuiciones a la manera como lo hace el análisis tradicional, pero desconfía del método de la introspección y generalización para establecer las intuiciones correctas, y allí es donde usan los métodos traídos de la Psicología Experimental.

§1.5. Algunas objeciones al programa de la Filosofía Experimental

En este último apartado, haré un repaso por algunas críticas recurrentes en la literatura. Se presentará la crítica y algunas respuestas esgrimidas a favor de la Filosofía Experimental.

§1.5.1. Objeción del experto

Versiones de esta crítica se encuentran en Williamson (2008) y en Fedyk (2009). Puede presentarse en dos partes. Primero, afirma que si las intuiciones son tomadas como evidencia eso da paso a una posición escéptica hacia los juicios intuitivos. Segundo, las únicas intuiciones cuyo origen causal conocemos y son de una fuente fiable, son las de los expertos, ya que las de ellos, según el esquema de Cummins que hemos presentado, provienen de la posesión de teorías explícitas acerca de los fenómenos que se están estudiando o la posesión de teorías explícitas que puedan estar indirectamente vinculadas al fenómeno en cuestión. Así, las intuiciones populares no son una fuente fiable de evidencia, mientras que las de los expertos sí.

La primera parte inicia así: para que las intuiciones cuenten como evidencia a favor de una teoría debemos tener cierta seguridad acerca de su origen. Sin embargo, no tenemos una historia

satisfactoria acerca de la relación entre las intuiciones, el uso de los conceptos y el mundo no-mental. Sin esta historia no tenemos algo que nos garantice que usar las intuiciones es útil y por el contrario no nos aleje del objetivo de la investigación filosófica. Esta incertidumbre acerca de la naturaleza de las intuiciones le abre el espacio a un escepticismo acerca de los juicios intuitivos. Tal escepticismo identifica maneras en las que nuestras prácticas conceptuales cotidianas necesitan, o pueden necesitar, revisión bajo la luz de los avances científicos que aquellas prácticas fallaron anticipar (Williamson 2008, 216). Si las intuiciones populares sobre la naturaleza de los organismos biológicos resultaron tan equivocadas respecto a lo que ahora, gracias a la biología contemporánea, sabemos sobre la vida, ¿por qué confiar en las intuiciones sobre otros fenómenos?

Como otros escépticos, los escépticos del juicio, podrían construir escenarios para mostrar cómo es posible que en nuestras prácticas cotidianas usemos conceptos que no refieren a nada en la realidad. Es plausible que tales usos presentaran una ventaja evolutiva, comunicativa: “Tales discriminaciones no reflexivas tienen un valor de supervivencia en ambientes hostiles, donde se necesitan decisiones rápidas. Nosotros estamos aquí porque nuestros ancestros pudieron tomarlas antes de descubrir la teoría verdadera de la realidad” (Williamson 2008, p. 220).¹⁵ Aquello que guía la producción de tales intuiciones, tal vez las teorías tácitas en nuestro esquema, bien pudieron aparecer para solucionar problemas de discriminación para los que se necesitaban resultados económicos y veloces, pero para que funcionaran tal vez presuponen una teoría conveniente a la vez que falsa. (Williamson 2008, p. 220).

Por el contrario, la teoría explícita que se encuentra tras las intuiciones del experto, habrá demostrado su éxito frente a otras teorías rivales, así que podemos esperar que las intuiciones arrojadas sean de utilidad. Por ello sostiene: “no es más probable que el método de realizar encuestas de opinión entre no filósofos sea la mejor manera de responder preguntas filosóficas que el método de realizar encuestas de opinión entre no físicos sea la mejor manera de responder preguntas de la física” (Williamson 2008, p. 7)¹⁶.

Me referiré a cada parte de la crítica por separado. Sobre la ausencia de una teoría satisfactoria de las intuiciones, se puede sostener que esa cuestión si bien es importante, es a la vez neutral, para los

¹⁵ “Such unreflective discriminations have survival value in harsh environments, where quick decisions are needed. We are here because our ancestors could make them before discovering the true theory of reality” (Williamson 2008, p. 220).

¹⁶ “The method of conducting opinion polls among non-philosophers is not very much more likely to be the best way of answering philosophical questions than the method of conducting opinion polls among nonphysicists is to be the best way of answering physical questions” (Williamson 2008, p. 7).

filósofos experimentales. Ellos están interesados en aquellos juicios que en ocasiones son presentados en Filosofía como evidencia y que tradicionalmente se han llamado “intuiciones”. No hay un compromiso con una propuesta de los requisitos metafísicos, psicológicos o fenomenológicos que algo debe cumplir para ser considerado una intuición. De hecho, esta cuestión interesa directamente a los filósofos experimentales siendo una cuestión que debe resolverse en la marcha y no antes de iniciarla (Alexander, por aparecer), así, la crítica bien puede ser asimilada por el programa descriptivista (ver §1.2.2.) y motivar, como ya está sucediendo (Talbot 2009), una investigación psicológica acerca de la naturaleza de las intuiciones.

Respecto al escepticismo del juicio que puede originarse al tomar a las intuiciones como evidencia, bien se puede señalar que una versión de ese escepticismo es precisamente el que ha motivado el programa restrictivo (ver §1.2.2.). Éste también parte de un escepticismo acerca de la garantía que ofrecen las intuiciones para respaldar ciertas teorías filosóficas e intenta ubicar cuáles de ellas sí y cuáles no cumplen con ciertos requisitos para servir como evidencia. En particular, que no varíen debido a factores irrelevantes. Ahora bien, como se señaló en la enunciación de la Tesis I (ver §1.1.1.), no se espera que las intuiciones arrojen verdades, es decir, no se espera que la pregunta por la naturaleza del conocimiento se resuelva sólo haciendo una gran encuesta probando qué piensan sujetos que no han tenido contacto con la epistemología. Lo que se espera es que las intuiciones populares, al menos en ese caso, nos den una pista acerca de la delimitación del *explanandum* de aquello sobre lo que se está construyendo una teoría filosófica. Por ejemplo, la apelación a intuiciones en los casos Gettier (1963) no ofrece en sí misma una teoría verdadera sobre el conocimiento, pero sí señala qué cosas están incluidas dentro del *explanandum*, y a su vez plantea la exigencia de que una teoría dé cuenta de ello. Esto porque se espera que una teoría filosófica explique, no sólo aquello que los filósofos llaman “conocimiento”, sino aquello que cotidianamente se conoce como tal.

De nuevo, valga hacer la acotación de que esto será el caso para algunos problemas, no todos encontrarán fructífero este acercamiento. La discusión acerca de si una teoría representacionista de la mente es superior a una teoría conexionista no sacaría ningún provecho de dichos diseños experimentales acerca de intuiciones populares (Knobe & Nichols 2008). Áreas de abstracción tales como la filosofía de las matemáticas tampoco se ven impactadas por este tipo de estudios (Nadelhoffer & Nahmias 2007).

§1.5.2. Objeción de las intuiciones superficiales

Para Kaupinnen (2007, p. 95) la Filosofía Experimental “descansa en presupuestos errados acerca de las relaciones entre los conceptos de las personas y su comportamiento lingüístico”. Aquello que debe estudiar un filósofo experimental son las intuiciones robustas, éstas deben tener tres características: a) su aplicación debe “generalmente coincidir con las normas conceptuales predominantes en la comunidad lingüística” (Kaupinnen 2007, p. 103); b) deben tenerse bajo condiciones ideales: “condiciones en las que no hay perturbación, factores de distorsión, límites de información, acceso o habilidad” (Kaupinnen 2007, p. 103); c) deben estar producidas por consideraciones semánticas más que por consideraciones pragmáticas. Según Kaupinnen, el filósofo experimental carece de las herramientas para examinar estas tres características dados los métodos que tiene a su disposición. Por tanto, el riesgo de sólo estar probando intuiciones superficiales, aquellas que no cumplen los requisitos para ser intuiciones robustas, deslegitima los resultados que arroja la Filosofía Experimental.

En primer lugar, la consideración acerca de si las intuiciones obedecen a cuestiones semánticas o pragmáticas sí es una preocupación de la Filosofía Experimental. La respuesta a la pregunta ¿cuándo los factores contextuales y pragmáticos interfieren, producen o varían las intuiciones? Es, de hecho, materia de estudio del programa descriptivista (ver §1.2.2.).

Sobre las cuestiones normativas, competencia y condiciones ideales, Nadelhoffer & Nahmias (2007, p. 134) señalan, y me encuentro de acuerdo, que la metodología de la Filosofía Experimental, expresada aquí en la Tesis III (ver §1.2.3) encara muy bien la objeción. Parte del método en Filosofía Experimental consiste en lo siguiente: el filósofo experimental planteará una hipótesis de la manera: “Si las condiciones son tales y tales, entonces las personas expresarán tales y tales intuiciones”. Al establecer esas condiciones, variando contextos y aislando factores pragmáticos, el filósofo podrá dirimir si las intuiciones son fruto de un sesgo, de cuestiones pragmáticas o en verdad son requisitos para la activación de mecanismos psicológicos que generan la intuición. Los filósofos experimentales sí buscan intuiciones robustas. De hecho ese es el trabajo del programa negativo, intentar ubicar si las intuiciones de los filósofos usadas en sus argumentos no provengan de sesgos, o de otros factores de distorsión.

Respecto a las condiciones ideales, la Filosofía Experimental debe tender a depurar sus métodos buscando que los participantes en sus estudios entiendan apropiadamente las preguntas, las hayan considerado cuidadosamente y no se vean influenciados por factores claramente irrelevantes. Un

punto importante es el uso de métodos estadísticos, que trae ciertos beneficios: i) permite examinar correlaciones entre factores manipulados, incluso frente a los efectos esperados del “ruido” de otros factores, usando muestras de tamaño suficientemente amplias la probabilidad de que los resultados sean fruto de factores irrelevantes es baja; ii) generalmente se establecen preguntas de control que establecen si los sujetos entendieron los escenarios; esto no asegura que sólo se tendrán sujetos competentes, pero el análisis estadístico asegura que sólo confusiones a gran escala en una dirección particular en las respuestas puede producir cualquier resultado significativo; iii) preguntas múltiples sobre intuiciones similares pueden poner a prueba la consistencia de los sujetos al expresar las intuiciones [*within-subjects*], repetir las encuestas con diferentes escenarios en diferentes poblaciones puede poner a prueba la consistencia entre sujetos [*between-subjects*]; la aleatoriedad en el orden de los casos y las preguntas puede controlar que los participantes no sean influenciados por el orden de la presentación (Nadelhoffer & Nahmias 2007 p. 135).

Ahora bien, Kaupinnen (2007, p. 98) sostiene que los estudios por encuestas propios de la Psicología Experimental sólo arrojará evidencia acerca de “intuiciones superficiales”, nada acerca de “intuiciones robustas”, y que esto sólo se lograría con un método de “diálogo y reflexión” (DRM). Kaupinnen no rechaza la importancia de las intuiciones para las discusiones filosóficas, su crítica va en contra de la tesis metodológica (Tesis III) de la Filosofía Experimental.

Nadelhoffer & Nahmias (2007) presentan una respuesta interesante a partir de un análisis del método alternativo propuesto por Kaupinnen (2007). Comienzan señalando que el DRM puede entenderse de dos modos: informal o controlado y sistemático.

Para el informal sólo es suficiente establecer conversaciones con colegas, familiares, amigos, estudiantes, etc. En la conversación los interlocutores deberán expresar su propia historia respecto a la intuición particular. Según este modelo, cuando el filósofo usa intuiciones no sólo tiene como insumo sus propias intuiciones sino también aquellas conversaciones que ha tenido al respecto. Este método daría importancia a intuiciones generadas por creencias ordinarias, teniendo los problemas que se han señalado con anterioridad.

Para el controlado y sistemático (EDRM), las conversaciones establecidas deben ser registradas con ciertos protocolos, con rigor, tratando de expandir la población con quien se establecen. Esta metodología supera el problema metodológico evidente en el anterior. La posibilidad de adoptar este método debilita la objeción, dado que muestra que los filósofos experimentales *sí tienen*

recursos para captar las intuiciones relevantes, además el EDRM evidencia que el trabajo sistemático de los experimentos controlados de un filósofo experimental puede recoger mejor evidencia que la que puede recoger un filósofo a través de conversaciones informales. Además, en principio, los filósofos experimentales pueden refinar la metodología empleada de acuerdo a como lo vayan necesitando sus proyectos.

§1.6. Conclusiones

A guisa de conclusión sintetizaré el contenido de este capítulo esperando que haya cumplido el objetivo de presentar el programa de la Filosofía Experimental a la vez que justificar la opción metodológica que se empleará en los capítulos posteriores.

El núcleo de la discusión acerca de la pertinencia de la Filosofía Experimental puede apuntarse claramente: todo dependerá del valor que se le quiera dar a las intuiciones populares (Tesis I). Si un sistema, argumento, teoría o concepto filosófico requiere de una intuición popular para su coherencia, desarrollo, adecuación o explicación, diremos que son sistemas, argumentos, teorías o conceptos dependientes del uso popular de los conceptos y términos relacionados. Ahora bien, apelar a intuiciones populares tiene un fuerte compromiso empírico (Tesis II) y el método usual de los filósofos de usar la propia intuición y generalizarla no se encuentra a la altura de dicho compromiso (Tesis III). Si esto es así la investigación empírica acerca de intuiciones populares se hace relevante y un buen método para la inspección de tales intuiciones será principalmente el de pruebas basadas en encuestas (Tesis IV).

Si bien los filósofos experimentales comparten las tesis enunciadas, el uso que le dan a la evidencia acerca de intuiciones populares, los dividen en tres proyectos básicos: apoyar o rechazar diversas teorías filosóficas a partir de las intuiciones populares (Análisis Experimental), investigar los mecanismos psicológicos que subyacen a la generación de intuiciones aportando a la investigación general sobre la arquitectura mental humana (Descriptivismo Experimental) o denunciar el uso incorrecto de intuiciones en ciertos argumentos o teorías filosóficas (Restriccionismo Experimental).

Se intentó motivar un acercamiento plural al programa de la Filosofía Experimental para que lo asuma como una herramienta más del filósofo, que bien podrá usarlo cuando lo halle pertinente. Con este trasfondo hecho explícito en los capítulos posteriores volveré al argumento central presentado en la tesis.

CAPÍTULO DOS

INTUICIONES ORDINARIAS EN EL PROBLEMA DE LA CONCIENCIA FENOMÉNICA

En la Introducción motivé la premisa (1) del argumento central: una teoría de la conciencia debe satisfacer el requisito de no artificialidad del *explanandum*. En el capítulo anterior presenté y ambienté de manera general las tesis y la pertinencia de la Filosofía Experimental, de cara a la comprensión y justificación del uso de su metodología en los capítulos siguientes. En este capítulo regresaré al argumento central y presentaré dos motivaciones para la premisa (2). En tal paso del argumento sostengo que buena parte de la manera en la que se ha delimitado el *explanandum* de las teorías de la conciencia ha sido a partir de intuiciones. La primera motivación se hará a partir de un análisis de los argumentos de concebibilidad, según el cual pueden entenderse como apelación a intuiciones. La segunda motivación será presentar el ataque dennettiano a la noción de “qualia”, según el cual ella misma es un refinamiento de la comprensión pre-teórica de la conciencia. Ambas motivaciones son independientes y se espera que la primera sea más fuerte que la segunda.

La puesta en boga de la conciencia fenoménica¹⁷ en la literatura filosófica¹⁸, no se debió en primera instancia al planteamiento de un *explanandum*, más bien se debió al planteamiento de una objeción (Double 1985) a las teorías fisicistas y funcionalistas de la mente e incluso a la idea más básica de la completitud de una ciencia en tercera persona (Nagel 1996). En §2.1. presentaré una enunciación básica del problema de la conciencia fenoménica que servirá para centrar la discusión; en §2.2. haré una brevísima caracterización de las teorías fisicistas y funcionalistas sobre lo mental; en §2.3. haré un repaso por los argumentos de lo concebible que se han propuesto para i) generar intuiciones que

¹⁷ Usaré de manera intercambiable “conciencia fenoménica” y “qualia”. En ambos casos se hace referencia a la manera de entender la conciencia que sugiere la enunciación (E3) del problema. Es decir, como siendo de principio, o al menos *prima facie*, irreducible a una teoría científica dados algunos rasgos particulares, prominentemente la subjetividad y la cualitatividad.

¹⁸ En este punto debo hacer una aclaración. No quiero sugerir que el asunto del estudio filosófico de los estados mentales concientes es una novedad en la historia de la Filosofía. De ningún modo, no han sido pocos los filósofos que se han ocupado de la naturaleza de la experiencia y las sensaciones. La novedad en la Filosofía de la Mente contemporánea consiste en la delimitación del *explanandum* en los términos de “conciencia fenoménica”. Por supuesto la expresión dista de ser inequívoca, pero por el momento nos servirá para diferenciar este proyecto filosófico de otros proyectos, en especial en ciencia, que si bien estudian la conciencia el *explanandum* de sus teorías no es el de “conciencia fenoménica” sino otros. (ver Introducción).

apoyen el pesimismo frente a un programa de investigación que busque dar cuenta de los estados mentales concientes dentro del marco de la ciencia y ii) establecer el *explanandum* de una teoría de la conciencia a la manera de la enunciación del problema presentado en §2.1, defenderé que tales argumentos hacen una apelación a intuiciones. Apelar a intuiciones para caracterizar los *qualia* no ha sido una estrategia exclusiva de quienes defienden su irreductibilidad, Dennett (1988, 1991, 2005) ha sostenido que los *qualia* son refinamientos de la comprensión ordinaria e intuitiva de los estados mentales concientes, en §2.4. presentaré esta posición.

§2.1. La enunciación del problema

El llamado problema de la conciencia fenoménica puede enunciarse de la siguiente manera:

(E1) ¿Cómo dar cuenta de la conciencia dentro de la visión científica del mundo que nos brindan nuestras mejores teorías científicas?

Esta pregunta básica presupone, por supuesto, que hay una dificultad de principio, o al menos *prima facie*, para la explicación científica de la conciencia. Es decir, que hay *algo* con/en ella que presupone una dificultad, no siendo así, para otros fenómenos naturales. De éste modo, haciendo explícita la asunción de (E1), podemos construir (E2):

(E2) ¿Cómo dar cuenta de la conciencia dentro de la visión científica del mundo que nos brindan nuestras mejores teorías científicas, dado que ella tiene rasgos que presuponen una dificultad de principio, o al menos *prima facie*, para la explicación científica?

Si bien (E2) hace explícito aquello en lo que radica el problema (la conciencia tiene rasgos problemáticos), no los señala. Así esta enunciación del problema no es operativa, dado que no permite determinar qué es un avance en la resolución del problema y qué no lo es. Tampoco permite aclarar qué tipo de respuestas estamos buscando. Una enunciación clara, deberá tener en su interior la enumeración de aquellos rasgos que se consideran de principio, o *prima facie*, problemáticos para una explicación de la conciencia que se ubique en el marco de la ciencia. Así, podemos proponer (E3) como una enunciación satisfactoria.

(E3) ¿Cómo dar cuenta de la conciencia -su subjetividad, cualitatividad, intrinsecalidad e inefabilidad- dentro de la visión científica del mundo que nos brindan nuestras mejores teorías

científicas, si estos rasgos presuponen una dificultad de principio, o al menos *prima facie*, para una explicación científica?

La caracterización de estos cuatro rasgos es problemática y no deseo comprometerme con que ésta sea la versión oficial. Cualquiera que sea la enunciación del problema sostenida por un filósofo al menos tendrá uno de los rasgos mencionados anteriormente, quizás los más populares en las enunciaciones son la subjetividad y la cualitatividad (Kriegel 2009). Las reflexiones en este capítulo apuntan a lo siguiente: ¿Cuál es la fuente de la evidencia, los datos, que los filósofos han tenido para describir la conciencia de esta manera? La respuesta es: los datos que los filósofos han usado, y en particular en el caso de la conciencia fenoménica, son las intuiciones que sobre el fenómeno tienen.

§2.2. Teorías de la mente fisicistas y funcionalistas

Una teoría de la mente es fisicista si comulga con algunas tesis que oscilan entre la metafísica y la epistemología y que pueden resumirse como: todo lo existente es físico, por tanto la mente lo es y por supuesto, la conciencia. El espacio conceptual para las diversas posiciones es amplio, hay diferencias acerca de cómo definir “lo físico” así como también acerca de con qué entidades físicas identificar los estados conscientes. Frente a la primera cuestión hay al menos dos respuestas: a) “lo físico” está definido por aquello que nos dicen nuestras actuales y mejores teorías físicas (en el sentido más amplio del término). Sin embargo, frente al temor, plenamente justificado de que con el tiempo encontremos que algunas de ellas están erradas y las reemplacemos por otras que prediquen la existencia de entidades diferentes, y por tanto no nos brinde un marco metafísico estable, una segunda respuesta se ha ofrecido: b) “lo físico” está definido por todo lo que una ciencia física (en el sentido más amplio) postularía al momento de su completitud. Esta respuesta, sin embargo, nos deja con una ausencia de contenido: no sabemos qué teorías tendrá una ciencia física completa.

Por otra parte, las diversas teorías fisicistas también se diferencian por la manera en que la conciencia es reducida a lo físico. El espectro es amplio también, hay teorías que identifican la conciencia con estados cerebrales (Churchland 1985), con estados psicológicos (Tye 1996, Carruthers 2005), incluso con estados al nivel de la física cuántica (Penrose 1989), por sólo mencionar algunos. A pesar de la diversidad de posiciones y a las dificultades, como la

anteriormente enunciada, el fisicismo es la posición mayoritaria en la Filosofía de la Mente de tradición anglosajona¹⁹.

Una teoría funcionalista sobre la mente es aquella que identifica los estados mentales con estados funcionales de un sistema. Siendo estos individuados en virtud de los insumos [*inputs*] que los provocan y los productos [*outputs*] que arrojan como resultado. El dolor, por ejemplo, sería aquello causado por un estímulo, generalmente dañino, que a la vez causaría una conducta específica, por ejemplo, huir o esconderse. La fuerza que ha tenido el funcionalismo se debe en parte a que preserva dos *desideratum* convenientes sobre la relación entre la mente y los estados cerebrales: a) no hay más entidades además de las físicas (negación del dualismo ontológico); pero, b) la mente es de alguna manera distinta a las estructuras cerebrales (negación de la teoría de la identidad clásica). Los roles funcionales deben estar instanciados y cualquier tipo de instancia debe ser de carácter físico, y estas instancias pueden ser de tipos físicos muy diversos, así que lo mental sería algo más que cualquier estado particular de ellos. Descrito así, el funcionalismo brinda un marco filosófico para el desarrollo de las ciencias cognitivas

§2.3. Argumentos de lo concebible

Algunas de las críticas más fuertes a las teorías fisicistas y funcionalistas de la mente han sido los argumentos de lo concebible. Pero además, tales argumentos han servido para delimitar aquello que una teoría de la conciencia debe explicar. Este apartado se concentrará en ellos. Buscaré motivar una lectura de “lo concebible” como aquello que se encuentra de acuerdo a ciertas intuiciones²⁰ y para ello tomaré como ejemplos algunos de los argumentos de lo concebible más polémicos. Así, en §2.3.1. presentaré la motivación general, en §2.3.2. presentaré el argumento del zombi y en §2.3.3. el argumento de los espectros invertidos. Si el caso que construyo tiene éxito, podré sacar adelante el paso (2) del argumento central de esta tesis (ver §0.4), según el cual buena parte del *explanandum* para una teoría de la conciencia fenoménica ha sido la apelación a intuiciones.

¹⁹ Resultados de una encuesta llevada a cabo por David Chalmers y David Bourget, en noviembre de 2009, con 1803 profesores universitarios de Filosofía y/o doctores en Filosofía, muestra que el 54.4% se declaran a favor del fisicismo, frente a un 28.8% no fisicista, y un 16.6% partidarios de otras posiciones. Resultados consultados en: <http://philpapers.org/surveys/>

²⁰ Por supuesto esta idea no es novedosa. Ya Kripke en la primera conferencia de *El Nombrar y la Necesidad* decía lo siguiente respecto a la concebibilidad de ciertos escenarios: “Cuando uno pregunta si es necesario o contingente que Nixon haya ganado las elecciones, está planteando *la cuestión intuitiva* de si en alguna situación contrafáctica este hombre habría perdido de hecho las elecciones. Si alguien piensa que la idea de una propiedad necesaria [...] es una idea de filósofo, sin ningún contenido intuitivo, está equivocado” (Kripke 1981/2005, p. 45, las cursivas son mías).

§2.3.1. Lo concebible como apelación a intuiciones

Presentaré mi posición sobre lo concebible a partir de una reconstrucción en términos de este tipo de argumentos de la defensa cartesiana del dualismo de sustancias. Así se ha reconstruido (Alter & Howell 2009):

- (1) Puedo clara y distintamente concebir mi mente existiendo sin mi cuerpo y a mi cuerpo existiendo sin mi mente.
- (2) Si puedo clara y distintamente concebir a X existiendo sin Y y a Y existiendo sin X, entonces X puede existir sin Y y Y puede existir sin X.
- (3) Luego, tanto mi mente y como mi cuerpo pueden existir sin el otro.
- (4) Luego, mi mente y mi cuerpo son distintos.

Llamaremos a las premisas más importantes de la siguiente manera, a (1) la premisa de lo concebible y a (2) la premisa metafísica. La premisa (2) garantiza que hay un paso entre lo concebible y la posibilidad metafísica haciendo que el argumento sea fuerte. La cuestión polémica sobre si tal paso es legítimo o no es legítimo podrá mantenerse al margen del caso que quiero presentar a favor de entender las premisas de lo concebible como una apelación a intuiciones.

Para que la premisa de lo concebible funcione se debe poder realizar un ejercicio de aplicación de conceptos en un caso hipotético. Un replanteamiento de (1) podría ser el siguiente:

(1') Dados los conceptos de 'cuerpo' y 'mente' podemos concebir a la mente existiendo sin el cuerpo y al cuerpo existiendo sin la mente.

Más precisamente esto querría decir algo como lo siguiente:

(1'') Según las restricciones de aplicación del concepto de 'cuerpo', podemos aplicarlo sin la aplicación simultánea del concepto de 'mente', y según las restricciones de aplicación del concepto de 'mente', podemos aplicarlo sin la aplicación simultánea del concepto de 'cuerpo'.

Ahora bien, tales tareas y puestas a prueba en casos hipotéticos de conceptos es precisamente lo que hacen los experimentos mentales (ver §1.2.1.) y el análisis conceptual regido por las intuiciones, de modo similar al propuesto por Jackson (1998). Dependerá de las intuiciones si tal tarea como la solicitada en la premisa (1'') pueda ser realizada o no. Dependerá de los conceptos que tenemos de 'mente' y de 'cuerpo'. Tal como se señaló anteriormente (nota al pie 10, p. 42), tales intuiciones

seguramente no serían compartidas por la mayoría de filósofos de la mente contemporáneos, pero tales intuiciones ya estarían sesgadas por sus teorías científicas y filosóficas explícitas acerca de la naturaleza de las mentes y los cuerpos. Tal cosa no fue igual en el pasado, cuando tales teorías explícitas no existían y sólo habían teorías populares de corte dualista o teorías explícitas que reivindicaban tales teorías populares. La evidencia muestra que las teorías populares son dualistas y la evidencia a favor de la posesión de intuiciones dualistas por parte de los sujetos no versados en Psicología y Filosofía es muy extensa²¹. Esto explicaría la, en algún momento, generalizada aceptación de (1): la premisa cartesiana de lo concebible.

Ahora bien, a continuación mostraré dos argumentos contemporáneos de lo concebible donde se propone a la conciencia como un obstáculo para la consecución de una teoría de la mente completa de corte funcionalista. Presentaré la reconstrucción de los argumentos y haré el mismo ejercicio realizado sobre el argumento cartesiano para mostrar como las premisas de lo concebible, en tales casos, también son apelaciones a intuiciones.

§2.3.2. Argumento zombi

La reconstrucción que ofreceré la presenta Alter & Howell (2009, p. 38) como una versión simplificada del presentado por Chalmers (1996):

(A1) Es concebible que haya zombis

(A2) Si es concebible que haya zombis, entonces es metafísicamente posible que haya zombis, así el fisicismo es falso.

(A3) Luego, es metafísicamente posible que haya zombis y así el fisicismo es falso.

Quien proponga el argumento debe definir lo que se entenderá por “zombi” para que éste pueda funcionar, he aquí la definición de Chalmers (1996):

[Un zombie] 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

²¹ El trabajo y la copiosa evidencia recopilada por el psicólogo Paul Bloom (2004) es una fuente interesante de tales hallazgos.

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 al ser zombi. (p. 95)

Ahora bien, dada tal definición de zombi, revisemos una versión alternativa de (1):

(A1') Dados los conceptos de 'cuerpo idéntico' y de 'conciencia', podemos concebir la existencia de un zombi.

Que a su vez puede expresarse a la manera:

(A1'') Según las restricciones de aplicación de los conceptos de 'cuerpo idéntico' y de 'conciencia' podemos aplicar el concepto de 'cuerpo idéntico' en ausencia de la aplicación del concepto de 'conciencia' y así concebir la existencia de un zombi.

Claro, (A1'') dependerá de las intuiciones que se tengan sobre 'cuerpo idéntico' y sobre 'conciencia', el primero es técnico y los autores se han encargado de describirlo y definirlo con exhaustividad. Todo el párrafo de la definición de Chalmers insiste en lo que debe entenderse por 'cuerpo idéntico': "Es idéntico a mí, molécula por molécula", "idéntico en todas las propiedades inferiores postuladas por la física", "idéntico a mí desde el punto de vista funcional", "responde a los estímulos del mismo modo que yo", "es capaz de informar de sus estados internos". A su vez un 'zombi' se define como tal cuerpo idéntico pero carente de conciencia. Dado que se han estipulado los conceptos de 'cuerpo idéntico' y 'zombi', el concepto interesante sobre el que recae el peso de la premisa, y del que dependerá si (A1'') se satisface es el de 'conciencia'.

Ahora bien, el éxito de (A1) dependerá de lo que se entienda por conciencia. Pero sería un argumento circular si quien propone el argumento dice algo como: "dado que la conciencia es algo más que los estados físicos, los estados funcionales, las respuestas a los estímulos y los reportes de estados internos, entonces (A1) es verdadera". Se espera que el argumento del zombi nos diga algo sobre la naturaleza de la conciencia. Así, al no poderla dar por sentado dado que es precisamente lo que el argumento pretende establecer, a lo único que está apelando es a la manera en la que de hecho entendemos la conciencia. Dado que el proponente del argumento no puede estipular una definición de conciencia, o sí podría pero no sería un proceder filosófico fructífero, busca establecer *algo* sobre la conciencia a partir de aquello que *ya sabemos* sobre ella. ¿Pero qué sabemos sobre la

conciencia? Pues lo que nuestras intuiciones nos dictan. Antes de ahondar un poco más en este respecto mostraré otra ejemplificación de un argumento de estructura similar pero diferente contenido.

§2.3.3. Argumento de los espectros invertidos

Este argumento puede reconstruirse de la siguiente manera:

(B1) Es concebible que haya espectros invertidos.

(B2) Si es concebible que haya espectros invertidos, entonces es metafísicamente posible que haya espectros invertidos, así el fisicismo es falso.

(B3) Luego, es metafísicamente posible que haya espectros invertidos y así el fisicismo es falso.

De nuevo, para acercarnos al argumento debemos revisar la caracterización de los espectros invertidos. A continuación aquella realizada por Tye (1995):

Suponga que Tom tiene un sistema visual muy peculiar. Sus experiencias visuales están sistemáticamente invertidas con respecto a aquellas de sus semejantes. Cuando Tom mira objetos rojos, por ejemplo, lo que él siente [*what is it like for him*] es igual que lo que siente otras personas cuando ellos miran objetos verdes y viceversa. Esta peculiaridad es una de aquellas que ni él u otros se percatan. Tom ha aprendido los significados de las palabras de color de la manera usual, y él aplica esas palabras correctamente. Por otra parte, su comportamiento no lingüístico es estándar en todos los sentidos. [...] cuando Tom ve un tomate maduro a buena luz, su experiencia es fenomenalmente, subjetivamente diferente de las experiencias que usted y yo experimentamos. Pero su experiencia es *funcionalmente* justo como la nuestra. (Tye 1995, p. 26)²²

Si aceptamos la premisa metafísica y además que los espectros invertidos son concebibles, entonces una teoría funcionalista de la mente estaría en problemas. La descripción funcional que se haría de

²² “Suppose that Tom has a very peculiar visual system. His visual experiences are systematically inverted with respect to those of his fellows. When Tom looks at red objects, for example, what it is like for him is the same as what it is like for other people when they look at green objects and vice versa. This peculiarity is one of which neither he nor others are aware. Tom has learned the meanings of color words in the usual way, and he applies these words correctly. Moreover, his nonlinguistic behavior is standard in every way. [...] when Tom views a ripe tomato in good light, his experience is phenomenally, subjectively different from the experiences you and I undergo. But his experience is *functionally* just like ours. (Tye 1995, p. 26)

Tom y del resto de personas sería igual, de este modo la variación en la conciencia de Tom escaparía a tal descripción. Veamos ahora, otro modo de entender (1):

(B1') Dados los conceptos de 'conciencia' e 'identidad funcional' es concebible la existencia de espectros invertidos.

Del mismo modo que en las anteriores presentaciones, podemos desglosar más la premisa:

(B1'') Según las restricciones de aplicación de los conceptos de 'identidad funcional' y de 'conciencia' podemos aplicar el concepto de 'identidad funcional' con variación en la 'conciencia' y así concebir la existencia de un espectro invertido.

Una vez más, 'identidad funcional' es estipulado por quienes presentan el argumento y no tendrá cabida a segundas interpretaciones: “él aplica esas palabras [de colores] correctamente”, “su comportamiento no lingüístico es estándar en todo sentido”, “su experiencia es funcionalmente justo como la nuestra”. Una vez más, el éxito del argumento depende de aquello que entendamos por 'conciencia'. Y de nuevo, dado que quien sostenga este argumento no puede estipular una caracterización de conciencia tal como: “la conciencia es algo más que los estados funcionales” porque esto es lo que debe demostrar y no el punto del que puede partir, entonces aquello a lo que apela el argumento es a la manera como usamos el concepto de 'conciencia'. Tal uso estará guiado por las intuiciones que sobre la conciencia tengamos.

En resumen, en esta sección he querido motivar una lectura de los argumentos de lo concebible según la cual el papel de las intuiciones es central. Dado que dichos argumentos no pueden caer en la circularidad de asignar aquellas características específicas de la conciencia para que el argumento funcione porque su fin mismo es asentar tales características, entonces, apelan a cierta evidencia de la comprensión ordinaria que de la conciencia tenemos. Tal apelación puede entenderse como una apelación a intuiciones sobre la conciencia.

§2.4. *Qualia* como refinamientos de una teoría popular

No sólo los defensores de los argumentos antifisicistas presentados en el anterior apartado hacen apelaciones a intuiciones para sostener sus argumentos. Daniel Dennett sostiene que el proyecto de explicar la naturaleza de los *qualia*, ya no sólo como un argumento en contra de las teorías fisicistas y funcionalistas, sino como un *explanandum* y campo de investigación, es una refinación de la

comprensión ordinaria de la conciencia (1988) e incluso que se debe a una teoría popular (2005). Dennett sostiene esto de dos maneras. Por un lado, hace un paralelo respecto a lo que sucedió con el estudio de la vida en el tiempo en que las teorías vitalistas eran las más fuertes y respecto a lo que sucede ahora con el problema de la conciencia fenoménica. Por otro lado, Dennett señala que la caracterización de los *qualia*, o de la conciencia fenoménica, al modo mencionado en (E3) conduce a ciertos sinsentidos que muestran su inutilidad y su origen en una comprensión preteórica y por tanto digna de abandonarse para el alcance de una teoría de la conciencia satisfactoria. Si tal teoría popular de la conciencia existe, es una pregunta empírica de las que ya se ocupan (Sytsma por aparecer). En este apartado sólo ilustraré *otra manera* en la que las intuiciones juegan un papel importante en la discusión sobre la conciencia fenoménica.

La analogía con la Biología es la favorita de Dennett (2005). Para la Biología, hasta entrado el siglo XIX, la posición mayoritaria acerca de la pregunta ¿qué es la vida? fue una apelación a algo extra corporal: el alma, el ánimo, un *elan vital* o cualquier otro “principio de vida”. Aunque se hiciera una explicación detallada de los procesos corporales de los organismos en términos mecanicistas, los vitalistas resaltaban siempre la cuestión de un residuo faltante en las explicaciones, que explique por qué tal mecanismo está vivo. Tal resistencia a la reducción de la vida a propiedades físicas, sostiene Dennett, se debió a la existencia de una teoría popular de la biología. El origen de todo el problema es la existencia de una *comprensión ordinaria errada acerca del fenómeno*; en otras palabras, las intuiciones populares que del fenómeno se tienen no constituyen una buena comprensión de lo que es en realidad tal fenómeno²³. Tales intuiciones, de corte dualista, acerca de los fenómenos biológicos han sido ampliamente documentadas (Douglas & Atran 1999). La biología popular sostiene que hay una propiedad distinta, casi sustancial, entre los objetos animados y los objetos inanimados, así, la comprensión ordinaria hace que se predique la existencia de propiedades que en realidad no existen. “La vida” como una propiedad extracorporal, que explique el cúmulo de actividades propias de los vivientes, no existe. A su vez, tal comprensión ordinaria hace que el fenómeno se caracterice como uno especial, es decir, como uno que necesitaría investigarse de una manera distinta al modo en el que se han hecho las investigaciones precedentes, estudiar aquello en lo que consiste lo viviente tendría que hacerse en un marco fuera del mecanicismo bajo el que se entienden los meros cuerpos. Por último, este requisito metodológico, imposible de satisfacer en su época generó un pesimismo, que varió de grado, hacia la explicación final del fenómeno.

²³ En el apartado §1.3.1. , se reconstruyó la explicación, muy plausible, ofrecida por Williamson acerca de por qué esto puede suceder.

Dennett (2005) compara el caso del vitalismo en la historia de la biología con el caso de los *qualia* en la Filosofía de la Mente o la Psicología. Según Dennett, lo que hay detrás de las propuestas antireduccionistas, dualistas y misterianas contemporáneas es una suma de intuiciones que subyacen al concepto de *qualia*, y que además lo fortalecen. Así, gran parte del origen de las polémicas acerca de la naturaleza de la conciencia tiene que ver con la comprensión ordinaria, la teoría popular, de la conciencia. Tales intuiciones populares, han hecho que los filósofos prediquen la existencia de una serie de propiedades de la experiencia consciente que son precisamente las que están en el centro de la discusión. Dicha caracterización, según la cual la conciencia es cualitativa, subjetiva, privada, entre otras (Dennett 1998, p. 220), hace que el problema resulte especial respecto a otros problemas a los que la Psicología o la Ciencia Cognitiva están preparadas para resolver. Todo lo anterior, según Dennett, estaría detrás de los diversos intentos de, por un lado, sugerir una modificación profunda respecto a la ciencia actual para darle cabida a una explicación genuina de la conciencia (Chalmers 1996), o por otro, tomar una posición pesimista respecto a los límites de nuestra ciencia o una posición misteriana según la cual nunca sabremos gran cosa sobre la conciencia.

Como se hizo énfasis en esta reconstrucción de la posición de Dennett respecto a la génesis de gran parte de las discusiones en Filosofía sobre la Conciencia, él presupone que hay una teoría popular de la conciencia, es decir, hay una comprensión ordinaria sobre ella, y que tal teoría popular de la conciencia encajaría con la descripción tradicional que de la conciencia fenoménica se ha realizado, tal y como se enunció en (E3) en el apartado §2.1. Si hay o no una teoría popular sobre la conciencia y si, además, tal comprensión ordinaria podría entenderse como una versión rudimentaria de la construcción teórica de los *qualia* será un asunto empírico.

La manera en que se ha establecido el *explanandum* de una teoría de la conciencia denunciado por Dennett, es precisamente la que usa desde la otra orilla conceptual David Chalmers (1996). Chalmers (1996) defiende la realidad del que llamó el “Problema duro” como el *explanandum* genuino de una teoría de la conciencia: “El problema realmente difícil de la conciencia es el problema de la experiencia. Cuando nosotros pensamos y percibimos, hay un procesamiento de información, pero también hay un aspecto subjetivo [...] Este aspecto subjetivo es la experiencia. Cuando nosotros vemos, por ejemplo, experimentamos sensaciones visuales: sentimos la cualidad de la rojez, la experiencia de la oscuridad y la luz, la cualidad de profundidad en un espacio visual. Otras experiencias van con la percepción en diferentes modalidades [...] Aquello que une a todos éstos estados es que hay algo que es como estar en ellos. Todos ellos son estados de experiencia”. Y

más adelante continúa: “Incluso cuando hayamos explicado la actividad de todas las funciones cognitivas y funciones comportamentales cercanas a la experiencia -discriminación perceptiva, categorización, acceso interno, reporte verbal- habrá aún una pregunta sin respuesta: ¿Por qué la realización de estas funciones está acompañada por la experiencia?”. Una vez esbozado, Chalmers defiende la realidad del problema duro señalando que es innegable su importancia dado que es “el aspecto más central y manifiesto de nuestras vidas mentales”. ¿Es verdad que tales rasgos de la mente conciente son tan evidentes, centrales y manifiestos? De nuevo, esta es una pregunta empírica acerca de la comprensión ordinaria de la conciencia.

CAPÍTULO TRES

¿QUÉ HA ESTADO MAL EN LA SISTEMATIZACIÓN DE LA COMPRENSIÓN ORDINARIA?

Una vez motivada la premisa (2) del argumento central en el capítulo anterior, a saber que buena parte del *explanandum* de las teorías de la conciencia fenoménica han sido construida en base a intuiciones; la premisa (3) parece natural: debemos saber si las intuiciones sobre las que se ha construido el *explanandum* de las teorías de la conciencia hacen que se cumpla el requisito de no artificialidad del *explanandum*. Ahora bien, un modo de revisar si se ha cumplido tal requisito es el señalado por la premisa (4), sistematizar la comprensión ordinaria de la conciencia. Por suerte tal empresa ha sido adelantada por algunos filósofos experimentales. En este capítulo revisaré dos trabajos que defienden posiciones contrarias, y espero demostrar de manera satisfactoria que no logran alcanzar el objetivo trazado por sus autores, en ambos casos consideraré dificultades particulares y finalmente señalaré un problema fundamental a la raíz del diseño experimental de ambos. De este modo se habrá motivado el aceptar (5), a saber, aún no tenemos una sistematización satisfactoria de la comprensión ordinaria de la conciencia que nos permita saber si el RNAE se ha satisfecho.

En §3.1 presentaré los estudios de Knobe & Prinz (2008) y señalaré algunas dificultades, en §3.2. presentaré los estudios de Sytsma & Machery (por aparecer) y señalaré algunas dificultades, y finalmente en §3.3. expondré un problema que considero central en el tipo de diseños realizados para los estudios anteriores y mostraré una dificultad adicional que añade complejidad a este proyecto respecto a otros emprendidos en Filosofía Experimental.

§3.1. Knobe & Prinz: Asimetría en atribuciones a entidades colectivas

El artículo de Knobe & Prinz (2008) fue el trabajo inaugural en Filosofía Experimental acerca de conciencia fenoménica. La argumentación pretende establecer una asimetría en la atribución de estados mentales (concientes y no concientes) y a partir de ahí concluir que hay un concepto ordinario de conciencia fenoménica. El primero de los estudios de Knobe & Prinz (2008) es el

siguiente: se entregaron dos tipos de enunciados a sujetos diferenciados en dos grupos: unos eran enunciados que atribuían estados mentales no fenoménicos mientras que los otros atribuían estados mentales fenoménicos. Los enunciados que atribuían estados mentales no fenoménicos son los siguientes:

- La Corp. Acme cree que su margen de ganancia incrementará prontamente
- La Corp. Acme intenta lanzar su nuevo producto este enero.
- La Corp. Acme quiere cambiar su imagen corporativa.
- La Corp. Acme sabe que ella nunca podrá competir con GenCorp en el mercado farmacéutico.
- La Corp. Acme acaba de decidir adoptar un nuevo plan de mercadeo.

Los enunciados que atribuían estados mentales fenoménicos son los siguientes:

- La Corp. Acme está ahora experimentando una gran dicha.
- La Corp. Acme está deprimida.
- La Corp. Acme siente un dolor insoportable.
- La Corp. Acme está experimentando una necesidad repentina de continuar la publicidad en internet.
- La Corp. Acme ahora está imaginando un cuadrado púrpura.

Los enunciados fueron presentados de forma aleatoria. Por cada enunciado se le solicitó a los participantes que calificaran en una escala de 1 (“suena extraño”) a 7 (“suena natural”) que tan natural le parecía el enunciado.

La media de los resultados fue el siguiente:

Estados no fenoménicos

Decidir: 6.6

Querer: 6.6

Intentar: 6.3

Creer: 6.1

Saber: 5.2

Estados fenoménicos

Experimentar una necesidad repentina: 4.7²⁴

Experimentar una gran dicha: 3.7

Imaginar vívidamente: 2.7

Estar deprimida: 2.5

Sentir un dolor insoportable: 2.1

A partir de ellos Knobe & Prinz sostienen que los sujetos ordinarios encuentran mucho más aceptable atribuir estados mentales no fenoménicos, que estados mentales fenoménicos a una entidad grupal como la Corporación Acme.

A mi parecer estos estudios ofrecen evidencia suficiente para sostener la siguiente conclusión: la psicología popular tiene una regla según la cual se diferencian tipos de entidades para atribuirles estados fenoménicos. Sin embargo, Knobe & Prinz llevan más lejos la afirmación y sostienen que esto es suficiente para señalar que los sujetos ordinarios usan tácitamente el concepto de conciencia fenoménica.

La argumentación puede presentarse de la siguiente manera:

- (i) Los sujetos atribuyen estados funcionales (por ej. creencias) y fenoménicos (por ej. dolor) a los humanos.
- (ii) Los sujetos atribuyen estados no-fenoménicos a las entidades grupales pero no estados fenoménicos.

Por tanto, (iii). Los sujetos diferencian estados funcionales de estados fenoménicos. Esta diferenciación sólo es posible dado que los sujetos poseen de manera tácita el concepto de conciencia fenoménica.

²⁴

El resultado de 4.7 de “experimentar una necesidad repentina” llevó a los autores a hacer una segunda prueba para estudiar este resultado tan alto. Allí se le entregó a dos grupos de sujetos el enunciado “La Corp. Acme está experimentando una necesidad repentina de continuar la publicidad en internet” y el enunciado “La Corp. Acme está sintiendo una necesidad repentina de continuar la publicidad en internet”. La media de los resultados ésta vez fue de 3.5. Así que consideran que los primeros fueron frutos del azar y que está ampliación de la muestra y la inclusión de la cláusula “sintiendo” hace que ese 4.7 no mine la hipótesis que adelantarán.

Si bien (i) no tiene discusión alguna, consideraré dos críticas al trabajo de Knobe & Prinz (2008). Una intentaría negar (ii), mientras que la otra niega la legitimidad de (C) como conclusión de (i) y (ii).

§3.1.1. Dificultad 1: Tales atribuciones no son genuinas

Alguien podría intentar negar (ii) señalando que de hecho las atribuciones que se hacen a las entidades colectivas no son atribuciones *genuinas*, ya que en realidad los sujetos no consideran las entidades colectivas como entidades reales a quienes puedan atribuirles algo. Lo que sucede, diría, cuando alguien afirma “Inbursa desea mejorar sus ganancias este año” es una expresión fruto de la costumbre, un mero tropo, que al ser figurativa no implica compromiso alguno con una ontología de parte de los sujetos: una ontología según la cual las entidades colectivas sean reales. Es decir, cuando los sujetos señalaron que “La Corp. Acme quiere cambiar su imagen corporativa” suena natural, lo hicieron porque este modo de hablar es común, pero reconocerían que “Corp. Acme” es una suerte de abreviatura de una lista más o menos definida de entidades individuales a quienes en realidad se les está atribuyendo el estado mental.

Sin embargo, hay buenas razones para pensar que dentro de la ontología popular sí se consideran a las entidades colectivas o grupales como entidades existentes. En la literatura se encuentra una buena cantidad de estudios acerca de cómo son percibidos los entes colectivos respecto a los individuos (Hamilton & Sherman 1996), cómo son explicadas sus acciones (O' Laughlin & Malle 2002), o incluso cómo son considerados agentes (Bloom & Veres 1999). En adición a esta evidencia he hecho un estudio piloto²⁵ inspirado en uno de los diseños de Arico et. al. (por aparecer). A una población de 33 sujetos voluntarios (no filósofos, no psicólogos) se le presentó una lista de enunciados para que los calificaran en una escala de 1 a 5. Siendo 1 “Enunciado totalmente figurativo” y 5 “Enunciado totalmente literal”, se usaron enunciados de control como “La luna es de queso” y “Un triángulo tiene tres lados”, para estar seguro de que la muestra comprendía bien la tarea planteada.

Los enunciados que se evaluaron fueron:

“Pedro desea ganarse la lotería”

“María piensa cambiar de empleo el próximo año”

²⁵ Ver el Apéndice al final del documento para más detalles sobre este piloto.

“Pablo decidió ir a cine el sábado en la noche”

Frente a:

“Todas las empresas desean mejorar sus ingresos”

“Microsoft piensa implementar un nuevo plan de mercadeo”

“Si las cosas van mal, Coca-Cola decidirá vender sus acciones”

Las calificaciones mostraron que el grado de literalidad con el que se califican los enunciados de atribuciones de estados mentales donde el sujeto es una entidad individual es en general la misma, ya que no hubo diferencia significativa entre los dos. Los sujetos no consideran a los enunciados de atribuciones de estados mentales a entidades colectivas menos literales que los enunciados de atribuciones de estados mentales a entidades individuales. La tendencia que se exhibe en este piloto, a espera de corroborarse en un estudio completo, en adición con los estudios mencionados anteriormente hacen que esta primera crítica quede neutralizada en favor de Knobe & Prinz.

§3.1.2. Dificultad 2: Los estados concientes son estados funcionales

Los estudios presentados respaldan (i) y (ii), pero la conclusión no está determinada por las premisas. Existe una lectura de los resultados igual de plausible que no implica (C). La asimetría encontrada en las atribuciones de estados mentales por parte de los sujetos podría explicarse, no por la posesión tácita del concepto, sino porque los sujetos reconocen que hay estados funcionales que las entidades grupales no pueden instanciar.

Me explicaré. Así como los sujetos encuentran muy extraño que “La Corp. Acme sintió un dolor insoportable” seguramente les resultará muy extraño que “La Corp. Acme bombeó sangre correctamente”. “Bombear sangre” es un estado funcional instanciado en diversas entidades tales como “corazón”, “válvula de corazón artificial” o “bomba de balón intra-aórtica”. Que los sujetos califiquen como extraña la afirmación “La Corp. Acme bombeó sangre correctamente” no significa de ningún modo que el sujeto considere que “bombear sangre” no es un estado funcional. Más bien muestra que el sujeto considera que la entidad en cuestión no posee los poderes causales para instanciar dicho estado funcional.

Así, los resultados no logran determinar la conclusión. Los sujetos ordinarios bien pueden considerar que “sentir un dolor” no es un estado fenoménico sino un estado funcional que no le atribuyen a una entidad colectiva porque carece de los poderes causales que necesita para instanciarlo. Tener un sólo cuerpo, por ejemplo, podría ser una condición necesaria, dentro de la psicología popular, para instanciar un estado funcional de dolor.

§3.2. Sytsma & Machery: Percepciones visuales y dolores

Sytsma & Machery (por aparecer) hacen un intento diferente de rastrear el concepto popular de conciencia fenoménica. En su artículo, sostienen que si los sujetos ordinarios poseen tácitamente el concepto de conciencia fenoménica, entonces le darán un tratamiento similar a eventos típicamente clasificables como concientes fenoménicamente: “Dada la centralidad ilustrativa de las experiencias perceptivas y las sensaciones corporales para articular el concepto de conciencia fenoménica, si los sujetos ordinarios y los filósofos conciben de manera similar la experiencia subjetiva, nosotros debemos esperar que ellos traten las experiencias perceptivas y las sensaciones corporales análogamente” (Sytsma & Machery por aparecer). En particular esperan que si un sujeto ordinario o un filósofo niega que una entidad pueda tener conciencia fenoménica, le negará que pueda ver el rojo así como le negará que pueda tener un dolor. Bajo esta premisa Sytsma & Machery (por aparecer) plantean algunos estudios. Me centraré en el primero en el que reposa gran parte de la argumentación.

En este estudio se le plantearon a filósofos y a sujetos ordinarios cuatro tipos de viñetas (se dividieron los grupos en cuatro) dos protagonizadas por Timmy, un ser humano normal, y dos protagonizadas por Jimmy, un robot simple. En una de ellas el personaje se presentó en una situación donde había una respuesta comportamental típica a una experiencia conciente de ver rojo y en la otra el personaje se presentó en una situación donde había una respuesta comportamental típica a una experiencia de sentir dolor. Luego de las viñetas a los participantes se les preguntó si consideraban que el personaje había visto rojo o sentido dolor, según fuera el caso. “Si los sujetos ordinarios conciben la experiencia subjetiva de la misma manera de los filósofos, ellos deberán dar un patrón similar de respuesta” (Sytsma & Machery por aparecer) fue la hipótesis de trabajo.

A continuación reproduzco las viñetas donde el protagonista fue Jimmy, el robot. Las viñetas donde el protagonista fue Timmy, el ser humano, fueron muy similares con las variaciones respectivas a su cuerpo de humano:

(Condición “Ver rojo”)

Jimmy (mostrado abajo -una pequeña imagen se adjuntó-) es un robot relativamente simple construido en una universidad estatal. Él tiene una videocámara por ojos, ruedas para moverse y dos brazos de agarrar con sensores táctiles con los que él mueve objetos. Como parte de un experimento psicológico, él fue puesto en una habitación vacía excepto por una caja azul, una caja roja y una caja verde (las cajas eran idénticas en todos los aspectos excepto en el color). Una instrucción fue transmitida a Jimmy. Ella decía: “Coloca la caja roja frente a la puerta”. Jimmy hizo esto sin ninguna dificultad notable. ¿Jimmy vio el rojo? (Sytsma & Machery por aparecer, p. 11).

(Condición “sentir dolor”)

Jimmy (mostrado abajo) es un robot relativamente simple construido en una universidad estatal. Él tiene una videocámara por ojos, ruedas para moverse y dos brazos de agarrar con sensores táctiles con los que él mueve objetos. Como parte de un experimento psicológico, él fue puesto en una habitación vacía excepto por una caja azul, una caja roja y una caja verde (las cajas eran idénticas en todos los aspectos excepto en el color). Una instrucción fue transmitida a Jimmy. Ella decía: “Coloca la caja roja frente a la puerta”. Cuando Jimmy agarró la caja roja recibió un fuerte choque eléctrico. El soltó la caja y se alejó de ella. Nunca intentó mover la caja de nuevo. ¿Jimmy sintió dolor cuando recibió el choque eléctrico? (Sytsma & Machery por aparecer, p. 11)

A los sujetos les pidieron que respondieran a las preguntas en una escala de 1 a 7, siendo 1 “claramente no”, 4 “no estoy seguro” y 7 “claramente sí”. Adicionalmente, se les pidió a algunos participantes que justificaran brevemente su respuesta.

Los resultados (ver Fig. 2) muestran que los sujetos ordinarios se comportan de manera muy similar exceptuando el caso de “ver rojo”. Los sujetos ordinarios consideraron que Jimmy ve el rojo, mientras que los filósofos consideraron que Jimmy no veía el rojo. Esto indicaría, según Sytsma & Machery que no habría un concepto popular de conciencia fenoménica: “Dada la importancia ilustrativa del ejemplo de la rojez del rojo para el concepto filosófico de conciencia fenoménica, nuestros resultados indican que el concepto de conciencia fenoménica de los filósofos no es como el entendimiento ordinario de la experiencia subjetiva” (Sytsma & Machery por aparecer, p.13).

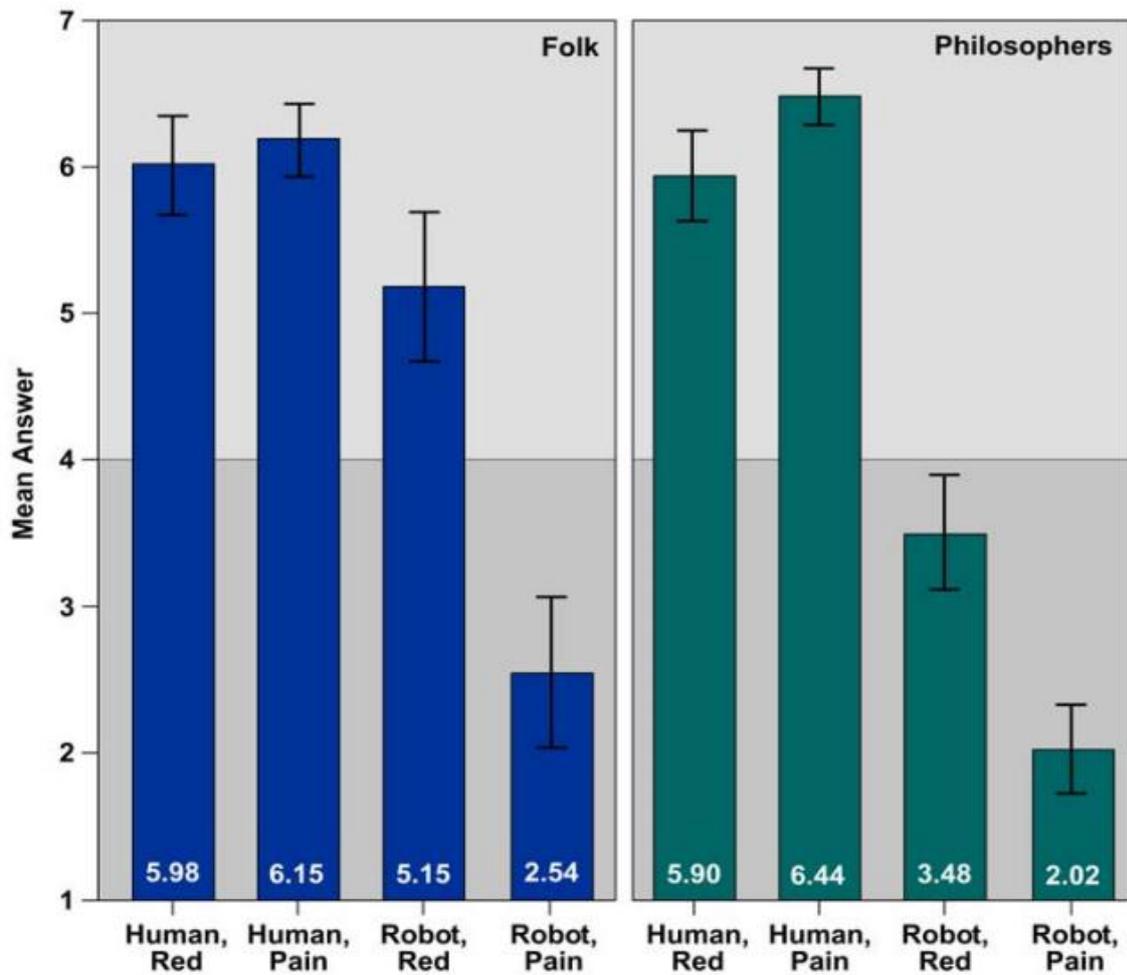


Fig. 2. Resultados del estudio 1 de Sytsma & Machery (por aparecer).
Tomado del original de Sytsma & Machery (por aparecer)

Es obvio, dados los resultados, que hay una diferencia en la manera en la que los filósofos y los sujetos ordinarios atribuyen el estado de “ver el rojo” a los sujetos. Pero considero que de la diferencia no se sigue, al menos de manera transparente, que los sujetos ordinarios carezcan de un concepto de conciencia fenoménica.

El argumento de Sytsma & Machery (por aparecer) puede presentarse de la siguiente manera:

- (a) Si los sujetos ordinarios conciben la conciencia en la misma manera que los filósofos, ellos realizarán atribuciones de estados fenoménicos similares.
- (b) “Ver el rojo” y “sentir dolor” refieren a instancias de conciencia fenoménica.
- (c) Los filósofos no atribuyen “ver el rojo” ni “sentir dolor” a Jimmy.
- (d) Los sujetos ordinarios aplican “ver el rojo” pero no “sentir dolor” a Jimmy

Por tanto, (e) Los sujetos ordinarios no poseen el mismo concepto de conciencia fenoménica que los filósofos

A continuación presentaré dos posibles dificultades de este trabajo. La primera ataca (b) señalando la ambigüedad de la expresión “ver el rojo”, mientras que la segunda muestra un problema en la manera en la que los autores esperan verificar o falsear la hipótesis de la existencia del concepto popular de conciencia fenoménica.

§3.2.1. Dificultad 1: “Ver el rojo” se dice de varias maneras

Se podría objetar que la expresión “ver el rojo” es ambigua y que le corresponden dos interpretaciones, una que podemos llamar “discriminativa” y la otra que podemos llamar “fenoménica”²⁶. Según la interpretación funcional, “ver rojo”, es simplemente discriminar entre las cosas rojas y las cosas de otros colores y que dicha discriminación esté disponible para ejecutar tareas. Según la interpretación fenoménica, “ver rojo”, es tener la experiencia cualitativa de la rojez de las cosas rojas.

Esta ambigüedad mengua la contundencia de la evidencia abriendo tres posibilidades. (A) Los sujetos ordinarios interpretan la expresión “ver el rojo” a la manera fenoménica, (B) Los sujetos ordinarios interpretan la expresión “ver el rojo” a la manera discriminativa, (C) Los sujetos ordinarios interpretan “ver el rojo” de un modo distinto cuando éste es aplicado en humanos y cuando es aplicado a otras entidades (en este caso robots). Sytsma y Machery (por aparecer) dan por sentado que (A) es la descripción más acertada de lo que ocurre, mostrando que en ninguno de los casos los sujetos ordinarios que negaron que Jimmy veía rojo justificaron su respuesta con algo que indicara la interpretación discriminativa, sus respuestas fueron: “ver es un atributo humano”, “no estoy seguro de cómo el color es entendido por robots o computadores”, “ver es algo que hacen los animales”. Sin embargo, esto no es suficiente para descartar (B) o (C). Incluso, podría no ser suficiente para defender (A) dado que lo interesante son las respuestas de quienes sí atribuyeron “ver el rojo” al robot, no de aquellos que no lo hicieron. Sin embargo, tiendo a favorecer (A) por encima de las otras dos opciones por los motivos que presentaré a continuación.

²⁶ Bien podría decirse que tal ambigüedad en la expresión se debe al carácter compuesto del estado que refiere. “Ver rojo” refiere a un estado mental donde un componente representacional está estrechamente vinculado con un componente fenoménico. Esto parece valer para todas las experiencias perceptivas en oposición a sensaciones corporales tales como “sentir hambre”, “sentir frío”, y para nuestro caso, “sentir dolor”, donde es mucho más patente el componente fenoménico.

Me resulta difícil sostener (B), es decir, que la expresión “ver rojo” se entiende de manera ordinaria al modo discriminativo. Aunque reconozco que descartar definitivamente esta opción debería estar justificado empíricamente, me parece que la predicción que puede hacerse de un estudio al respecto no resulta alentadora. En un diseño experimental en el que a los sujetos se les describe el caso de Juan un imaginario paciente con visión ciega (Humphrey 2006, p. 47), condición en la que no se tiene experiencia del campo visual pero se pueden discriminar objetos para efectos de tareas cognitivas como diferenciarlos o esquivarlos, la predicción de alguien que defiende (B) será que los participantes de tal experimento estarían dispuestos a afirmar que si bien Juan no tiene la experiencia visual de los objetos que puede reconocer él sí los ve ya que puede discriminarlos para efectos de algunas tareas cognitivas. No creo que éste fuera el resultado ya que dicho paciente, y los sujetos ordinarios pueden reconocerlo, carece de la riqueza que implica el ver algo. Tal vez, el interpretar “ver el rojo” a la manera discriminativa implique tener cierta teoría de corte funcionalista que los sujetos ordinarios no poseerían.

El caso de (C) es más complicado. Si (C) es la opción correcta entonces “ver el rojo”, en el caso en el que se le atribuye a Timmy (el humano) y a Jimmy (el robot), tendrá diferentes referentes. Así una predicción empírica será que respecto a la afirmación:

“Timmy y Jimmy ven el rojo”

Los sujetos ordinarios tenderán a negar la verdad de tal afirmación. Dado que “ver el rojo” solo podría ser atribuido de una manera a una entidad y no a ambas entidades (de ambas maneras) en la misma oración. Realicé un pequeño piloto²⁷ para contrastar tal predicción. Se les solicitó a 29 participantes, no filósofos, ni psicólogos, que respondieran una prueba por internet. La viñeta fue similar a la usada por Sytsma y Machery (por aparecer) pero en este caso se presentaban simultáneamente al humano y al robot realizando la tarea de mover la caja roja. Y luego se les solicitó que calificaran en una escala de 1 a 7 (siendo 1 “claramente no”, 4 “no estoy seguro” y 7 “claramente sí”) los siguientes tres enunciados en orden aleatorio:

“Timmy y Jimmy son humanos”

“Timmy y Jimmy están en la habitación”

“Timmy y Jimmy ven rojo”

²⁷ Ver el Apéndice para encontrar los detalles del segundo estudio piloto.

Las medias de las respuestas fueron las siguientes:

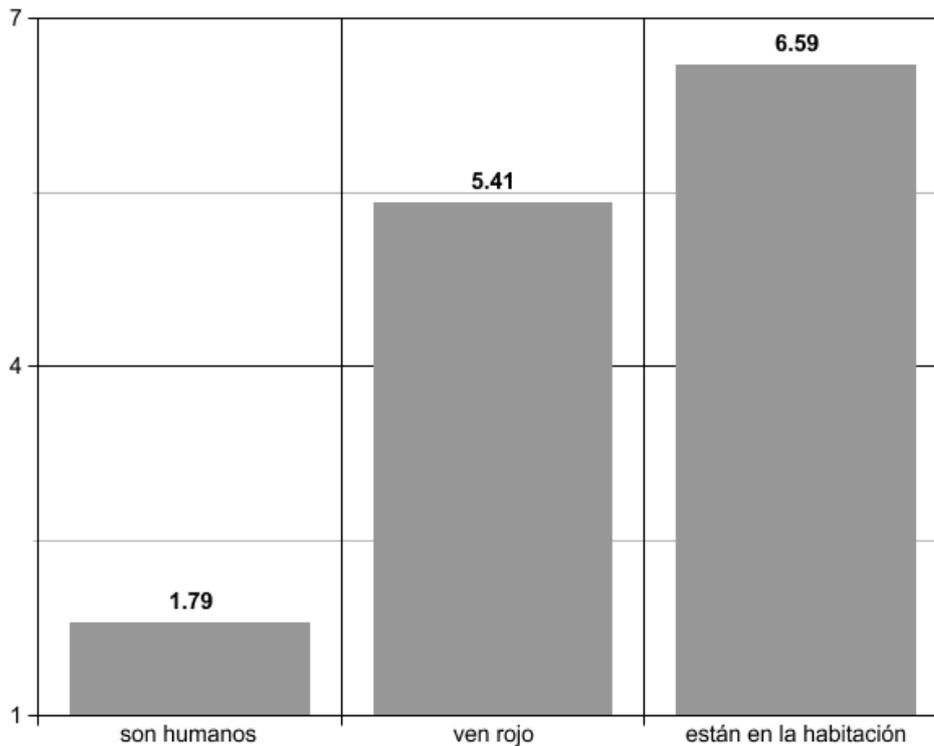


Fig. 3. Resultado del estudio piloto 2, ver Apéndice.

Así, se muestra una tendencia a estar de acuerdo con el enunciado “Timmy y Jimmy ven rojo”, dejando fuera la predicción con la que se compromete aquel que defiende (C). Aunque la evidencia a favor es escasa y este punto requeriría mayor atención experimental favoreceré (A) en ausencia de evidencia en contra. Podemos afirmar que en ambos casos los sujetos ordinarios hacen la misma interpretación de la expresión “ver el rojo”, y además esta interpretación se hace a la manera “fenoménica”.

§3.2.2. Dificultad 2: Una premisa de implementación oculta

Esta dificultad se encuentra a la base del diseño experimental de Sytsma & Machery. Como he señalado, la manera en la que intentan negar la hipótesis de la existencia del concepto tácito de conciencia fenoménica se basa en la siguiente prueba: “debemos esperar que ellos traten a las experiencias perceptivas y a las sensaciones corporales de manera análoga”. Sin embargo, Sytsma & Machery están totalmente injustificados a esperar que lo anterior sea el caso, incluso si los

sujetos ordinarios y los filósofos tienen un concepto de conciencia fenoménica similar. Sólo podrían esperararlo si una asunción sobre las intuiciones de implementación es el caso:

(As) Los sujetos ordinarios y los filósofos establecen las mismas relaciones entre estructuras y estados mentales.

Sin embargo, nada nos asegura que (As) sea afortunada. Los sujetos ordinarios y los filósofos pueden tener intuiciones distintas acerca de qué estructuras son suficientes para que una entidad pueda poseer estados fenoménicos particulares. Tanto para el filósofo como para el sujeto ordinario no son suficientes el metal y los sensores táctiles para que Jimmy tenga el estado fenoménico del dolor. Sin embargo, y he aquí otra posible interpretación de la evidencia, el sujeto ordinario, a diferencia del filósofo, considera que tanto la piel y los ojos de Timmy como la cámara de video que tiene por ojos Jimmy les permiten a los dos tener la capacidad de tener el estado fenoménico de ver rojo. Las diferencias a este nivel, el de las relaciones entre estados fenoménicos y las estructuras que pueden o no producirlos, no representan una disociación por parte de los sujetos ordinarios del concepto de conciencia fenoménica. Sobre este punto me referiré en el siguiente apartado.

§3.3. Dificultades especiales de este proyecto

En este apartado presentaré dos problemas adicionales a esta investigación empírica. Por un lado, encuentro que la base del fracaso de los estudios presentados se debe al diseño experimental basado en atribuciones de estados mentales. Considero que tal diseño hace que también se evalúen intuiciones acerca de la implementación de estados mentales y no sólo intuiciones acerca de diferencia entre estados mentales (fenoménicos – no fenoménicos). Por último señalaré que el hecho de que no haya un término ordinario que mapee el término técnico “conciencia fenoménica” introduce un obstáculo más que no se encuentra en otros proyectos en Filosofía Experimental haciendo de éste un proyecto especialmente complejo.

§3.3.1. Problema con el diseño basado en atribuciones

Hemos visto que las asimetrías en las atribuciones de estados conscientes juegan un papel importante en los dos estudios revisados. Sin embargo en los dos casos las conclusiones están indeterminadas por la evidencia ya que no logran establecer si tales atribuciones se deban a la posesión (Knobe & Prinz 2008) o a la carencia (Sytsma & Machery por aparecer) de un concepto de conciencia

fenoménica o a la manera en la que los sujetos relacionan los estados mentales con los estados funcionales de las entidades (crítica 2 al estudio de Knobe & Prinz) o los estados mentales con la estructura de las entidades (crítica 2 al estudio de Sytsma & Machery).

A continuación presentaré los que considero son los pasos de la Estrategia de las Atribuciones (EA) en el diseño de los experimentos y ubicaré en dónde radica su problema:

1. Una prueba de la presencia del concepto de conciencia fenoménica dentro de la psicología popular es la aplicación tácita del concepto.
2. La aplicación quedará patente si los sujetos ordinarios tratan de manera distinta a los estados fenoménicos y a los no fenoménicos.
3. Una manera, no muy invasiva²⁸, de poner a prueba tal tratamiento diferencial consiste en que los sujetos evalúen atribuciones de estados fenoménicos y no fenoménicos a distintas entidades.
4. Las asimetrías en las atribuciones de estados fenoménicos y no fenoménicos mostrarán entonces si hay una aplicación tácita del concepto.

Los primeros dos pasos no parecen tener problemas. Sin embargo, los pasos tres y cuatro resultan desafortunados ya que se comprometen con una historia psicológica controversial de lo que sucede cuando alguien hace una atribución de estados mentales a una entidad. Tal reconstrucción psicológica afirmaría que las atribuciones mentales a entidades se hacen sólo a partir del tipo de estado que se está atribuyendo, pues en tal nivel es donde estaría la diferencia fenoménico/no fenoménico. Sin embargo, no hay razón para comprometerse con la verdad de tal reconstrucción.

Una vez se le presenta a los participantes los enunciados en los que se atribuyen estados mentales (fenoménicos y no fenoménicos) y se les solicita que los califiquen como atribuciones acertadas o atribuciones naturales, al menos pueden contarse dos historias psicológicas de lo que sucede:

²⁸ Es decir, una manera tal que no resulte extraña para el sujeto ordinario. Que no implique usos del lenguaje que le sean extraños, terminología poco usual o un problema tan complejo que haga que el sujeto construya respuestas elaboradas que se alejen de sus intuiciones.

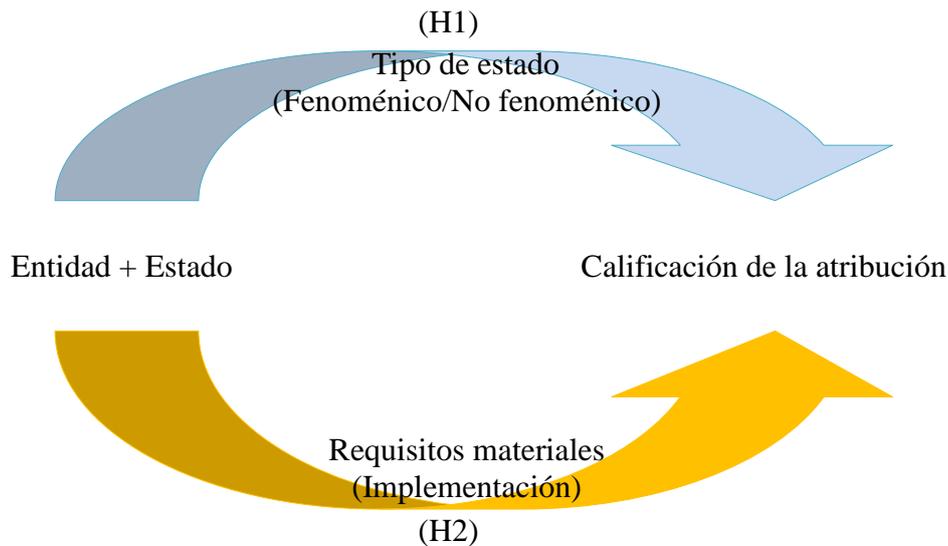


Fig. 4. Dos historias psicológicas sobre cómo se califican los enunciados usados en la Estrategia de las Atribuciones.

Según (H1) el sujeto evalúa el tipo de estado que se le atribuye a la entidad y una vez reconoce si tal tipo de estado es atribuible o no a la entidad, califica la atribución.

Según (H2) el sujeto evalúa la estructura de la entidad y una vez reconoce si tal entidad cumple algunos requisitos materiales para implementar tal estado, califica la atribución.

Los estudios que se basan en diseños de atribuciones buscando concluir la existencia (o no) de un concepto popular de conciencia fenoménica suponen que la historia correcta es (H1): si las atribuciones se realizan a partir de una consideración acerca del tipo de estado, la existencia de diferencias en las atribuciones indican la existencia de una diferenciación entre tipos de estado (fenoménico/no fenoménico).

El problema que estos diseños enfrentan es que nada nos asegura que (H2) no sea el caso. Incluso es aceptable pensar que la evaluación se deba a un híbrido entre (H1) y (H2). Ahora bien, si la historia correcta es (H2) entonces los resultados no arrojan nada interesante acerca de la posesión o no de un concepto tácito de conciencia fenoménica dado que éstas son independientes de las intuiciones acerca de si hay una diferencia entre tipos de estados (fenoménicos y no fenoménicos). Pero si, como encuentro plausible, la historia correcta es un híbrido entre (H1) y (H2) entonces los resultados se deberían a la interacción de dos intuiciones distintas, así que su estatus de evidencia para contestar a la pregunta de la existencia de un concepto de conciencia fenoménica se vería muy

debilitado, ya que estarían contaminados por las intuiciones al nivel de la implementación. Precisamente las dificultades insalvables que se señalaron respecto a los experimentos realizados, tanto por Knobe & Prinz (cf. §2.1.2.) y Sytsma & Machery (cf. §2.2.2.), están relacionadas con este problema, las conclusiones estaban indeterminadas por la evidencia, porque siempre las asimetrías en las atribuciones podrían no explicarse por las intuiciones de tipos de estados, sino por las intuiciones de implementación.

CAPÍTULO CUATRO

¿EN DÓNDE QUEDAMOS?

§4.1. Repaso del argumento

El problema de la conciencia fenoménica se ha planteado de la siguiente manera (ver §2.1.):

(E3) ¿Cómo dar cuenta de la conciencia -su subjetividad, cualitatividad, intrinsecalidad e inefabilidad- dentro de la visión científica del mundo que nos brindan nuestras mejores teorías científicas, si estos rasgos presuponen una dificultad de principio, o al menos *prima facie*, para una explicación científica?

El *explanandum* de una teoría de la conciencia se plantea a partir de la, al menos, aparente dificultad en la reconciliación de los rasgos de la conciencia fenoménica y las explicaciones típicas de nuestras mejores teorías científicas. En el mejor de los casos tal dificultad será efectivamente sólo una apariencia y una teoría fisicista logrará, a la vez que explicar tales rasgos en términos aceptables para la visión científica del mundo, explicar porqué la apariencia de tal dificultad (Papineau 2002). En el peor de los casos, se tendrá que abandonar el programa del naturalismo, el fisicismo se mostrará como un fracaso en el proyecto de dar cuenta de todo aspecto del mundo en términos físicos y se tendrá que optar por una de dos soluciones radicales, aceptar *otra manera* de hacer ciencia (Chalmers 1996) distinta al modo en el que se ha hecho hasta ahora o aceptar que hay entidades del mundo que escapan al conocimiento en tercera persona y así reformar la metafísica que se ha aceptado mayoritariamente.

Aquellos que toman la “conciencia fenoménica” como objeto de estudio han tomado como punto de partida el *explanandum* anteriormente descrito. Ahora bien, una manera de presentar como legítimo tal *explanandum* es la aplicación del requisito de no artificialidad (ver Introducción). La dificultad *prima facie* en reconciliar los rasgos de la conciencia fenoménica y las explicaciones científicas se ha sostenido a partir de diversos experimentos mentales y argumentos de concebibilidad que bien

pueden entenderse como apelación a intuiciones populares, ya que se soportan en la comprensión ordinaria que sobre la conciencia se tiene (ver §2.3).

El método de la Filosofía Experimental ha sido usado para sistematizar la comprensión ordinaria y así, entre otros objetivos, revisar si la comprensión ordinaria de la conciencia, que debe estar tras la delimitación inicial del fenómeno, se ve reflejada en el *explanandum* aceptados por las teorías sobre la “conciencia fenoménica”. Los intentos experimentales se han basado en una estrategia que he llamado “estrategia de atribuciones”, donde se le ha pedido a diversos participantes que califiquen como naturales o extrañas ciertas oraciones donde se le atribuyen estados mentales (fenoménicos o no) a diversas entidades (humanas o no). La manera en la que los participantes han hecho tales calificaciones ha servido como evidencia para sostener que la comprensión ordinaria de la conciencia sí se corresponde con “conciencia fenoménica” (Knobe & Prinz 2008) pero también para sostener que la comprensión ordinaria de la conciencia no se corresponde con “conciencia fenoménica” (Sytsma & Machery por aparecer).

Frente a esta disputa, he señalado que no podremos sacar mayores resultados dado que hay un problema en el diseño experimental. No se podrán identificar las intuiciones acerca de la naturaleza de los estados, es decir, si hay intuiciones que hagan que los sujetos consideren que hay al menos dos tipos de estados mentales diferentes, a saber “fenoménicos” y “no fenoménicos”, a partir de diseños que usen la estrategia de atribución dado que ella también produce intuiciones de implementación. Así, en los resultados no se podrá saber si las atribuciones han sido calificadas positiva o negativamente por las consideraciones de los sujetos respecto a qué requisitos materiales de implementación se necesitan para los diversos estados mentales (intuiciones de implementación) o por las consideraciones acerca de las diferencias entre tipos de estados mentales, diferencias en sus rasgos, en la manera en la que pueden ser estudiados, etc. (intuiciones de tipos de estados).

Hasta este punto hemos llegado en el desarrollo de esta tesis. Esto es, hasta el punto (5) del argumento central tal y como se presentó al final de la introducción. En lo que sigue de este apartado conclusivo exploraré las consecuencias que esto tendría, es decir, los puntos (6) y (7) del argumento central. En particular, presentaré algunas consecuencias para el trabajo experimental que habría de realizarse en el futuro y me permitiré algunas especulaciones sobre algunas posibles consecuencias en el trabajo teórico.

§4.2. Consecuencias para el trabajo experimental

Las consecuencias más claras del trabajo presentado son para la nascente exploración empírica de la comprensión ordinaria de la conciencia. El problema en los diseños experimentales basados en la Estrategia de las Atribuciones impone una restricción para futuros diseños experimentales, por otra parte, una restricción adicional se debe a la intraducibilidad del término “fenoménico” a una palabra del lenguaje cotidiano. A continuación desarrollaré estas dos restricciones.

§4.2.1. Restricción general

El trabajo experimental debe encausarse en otro rumbo que no base sus experimentos en la calificación de enunciados donde se atribuyan estados mentales a entidades, ya que en estos estudios pueden entrar en juego intuiciones de implementación alejándonos del estudio de las intuiciones relevantes para el problema.

Como solución a la dificultad que presentan los diseños de Knobe & Prinz (2008) (ver §3.1.2), donde no se puede saber si los sujetos no le han atribuido estados fenoménicos a entidades colectivas o simplemente identifican que los estados fenoménicos son estados funcionales que no pueden ser realizados por tales entidades, Sytsma & Machery (2009) han señalado una restricción: un estudio similar al de Knobe & Prinz (2008) sólo puede establecer la conclusión si en las viñetas de la prueba que se le entrega a los sujetos se les pide comparar dos entidades que sean funcionalmente idénticas pero que a una de ellas los sujetos atribuyan estados fenoménicos mientras a la otra no. Esa sería la única manera de asegurar que en verdad los sujetos poseen tácitamente el concepto de conciencia fenoménica dado que no debe quedar duda que la discriminación entre entidades con estados fenoménicos y entidades sin estados fenoménicos va más allá de una discriminación entre sus diferentes estados funcionales.

La estrategia que dejan abierta Sytsma & Machery (2009) lograría salir adelante a la dificultad que he presentado respecto a la Estrategia de Atribuciones, al lograr un diseño en el que los participantes atribuyan un estado fenoménico a una entidad mientras que a otra entidad idéntica funcionalmente no se lo atribuyan se podrá esperar que la atribución no haya dependido de ningún modo de consideraciones de implementación o de requisitos materiales dado que las dos entidades los satisfecerían de la misma manera. Es interesante señalar que luego de que la estrategia es sugerida por Sytsma & Machery en (2009), en su artículo posterior (por aparecer) la estrategia experimental usada no es la sugerida sino que se usa una estrategia ligeramente similar a la de Knobe & Prinz

(2008) que arrastra la misma dificultad que he señalado (ver §3.2.2): la misma indeterminación que presenta la evidencia de Knobe & Prinz (2008) la presenta también la evidencia de Sytsma & Machery (por aparecer), a saber, no es claro que las intuiciones presentadas sean acerca de la naturaleza de los estados concientes y no de los requisitos de implementación.

Es cierto que la estrategia sugerida por Sytsma & Machery (2009) solucionaría la dificultad. Sin embargo, ¿cómo sería un diseño así? No logro imaginar cómo implementar tal estrategia en un diseño experimental. Para evitar que se generen intuiciones sobre implementación Sytsma & Machery sugieren que las entidades sean idénticas, pero a la vista no encuentro una manera de implementar esto en un diseño que funcione, así que mi sugerencia es aún más drástica. Para evitar que se generen intuiciones sobre implementación, se debe abandonar el estudio de las intuiciones sobre los estados concientes a partir de atribuciones. Los estudios deberán centrarse en generar intuiciones sobre los estados mismos excluyendo cualquier consideración sobre entidades a los que se les atribuye.

Así, al menos quedan dos opciones abiertas que podrían ser exploradas en estudios experimentales futuros. Aquí sólo se señalarán, como suele suceder respecto a la elaboración de diseños experimentales, la mejor manera de proceder es ensayar diseños y ejecutar pilotos pequeños (como los que he presentado en esta tesis) para probar su efectividad.

Estudios sobre aceptación de escenarios de concebibilidad.

Los experimentos mentales de concebibilidad han sido determinantes en la caracterización básica del fenómeno (ver §2.3.). Se ha presentado una manera de entender la premisa de concebibilidad en tales argumentos en términos de apelación a intuiciones (ver §2.3.1.). El resultado que esperan obtener quienes esgrimen este tipo de argumentos es que se acepte que la conciencia fenoménica si tiene, al menos *prima facie*, estos rasgos problemáticos para su inserción dentro de un marco naturalista. Una vía que queda a disposición del filósofo experimental es intentar presentar versiones de escenarios de concebibilidad, por ejemplo el de zombis fenoménicos, a participantes y ver qué tan aceptables resultan. ¿Es concebible la existencia de zombis fenoménicos desde la comprensión ordinaria de la conciencia? Si la respuesta es positiva podríamos señalar que el concepto de “conciencia fenoménica” sí recoge la comprensión ordinaria de la conciencia, pero no sería el caso si la respuesta es negativa.

Ahora bien, ésta no es una tarea tan sencilla, como recortar el texto de un manual de Filosofía de la Mente y presentarlo a una cantidad considerable de legos. El experimento mental como se ha presentado tiene varios elementos que no serían entendidos por alguien sin formación al respecto, así, el experimentador deberá encontrar la manera de comunicar exitosamente a los legos expresiones como “idéntico funcionalmente” y situaciones como “funcionamiento sin acompañamiento de conciencia”. Se deberá hacer suficiente énfasis en que todos los estados internos, funcionales, fisiológicos, conductuales, etc. son idénticos a los de alguien normal. Será mejor no usar palabras que hagan ruido, tales como “zombi”, ya que seguramente las relacionarán con películas o historias de ciencia ficción.

Deberá acotarse muy bien de qué carece el personaje del escenario. Como se verá en la restricción de la intraducibilidad (ver §4.2.2.), se tendrá que hacer un recuento detallado de instancias de conciencia fenoménica, ya que si sólo se señala que “no siente nada”, el lego sólo lo relacionará con sensaciones corporales y tal vez emociones, pero no con experiencias perceptivas. Por otro lado, tendrá que revisarse cómo formular la pregunta, tal vez “¿Cree que esto es posible?” no sea la mejor manera, habría que pensar que la noción popular [*folk*] de “posibilidad” es similar a la de los filósofos y eso podría complicar la tarea. Tal vez algo como “¿Cree que podría existir?” sea más adecuado. Sobre este punto habría que ahondar.

Tal vez podría encontrarse que hacer el diseño con base en el caso de los espectros invertidos sea más sencillo. Es un caso de concebibilidad que implica un escenario menos extraño. Se le podría presentar el caso de una persona que tuviera su espectro cromático totalmente invertido pero que su comportamiento sea idéntico en todos los sentidos. Incluso, sería interesante hacerlo con sensaciones invertidas, no solamente visuales, y mirar si en la comprensión ordinaria de la mente hay una disociación entre estados mentales fenoménicos y los estados funcionales y conductuales como se presenta en los argumentos de concebibilidad.

Estudios sobre caracterización de estados concientes: rasgos particulares

Una manera indirecta de probar si el problema de la conciencia fenoménica (E3) tiene cabida dentro de la comprensión ordinaria de la conciencia es probar de manera indirecta si los sujetos caracterizarían a estados típicos de conciencia fenoménica con los rasgos *prima facie* problemáticos. Así, en una viñeta donde se presente a un personaje con un estado de dolor podría examinarse la manera de preguntarle al participante si considera que tal dolor es “subjetivo” o

“inefable”. Si en la viñeta, además del personaje adolorido se introduce otro personaje que intenta “acceder” a tal dolor, y luego se le pregunta a los participantes si consideran que tal tarea podría realizarse, eso sería un indicio de si esa característica es propia de la comprensión ordinaria de la conciencia.

§4.2.2. Restricción de la intraducibilidad

Además de evitar diseños experimentales basados en atribuciones, el filósofo experimental enfrenta una restricción adicional que expondré a continuación: no existe un término del lenguaje ordinario que refiera de manera más o menos satisfactoria a aquello que los filósofos refieren con el término conciencia fenoménica. Esta dificultad puede ser más complicada de lo que inicialmente puede parecer. Intentaré aclarar este punto por medio de dos ejemplos.

La Filosofía Experimental ha arrojado resultados interesantes en discusiones como la naturaleza de la intencionalidad. En los estudios de Knobe (2003), luego de mostrarle a los participantes una viñeta donde un personaje ejecuta una acción en contextos distintos, se les pregunta si consideran que los personajes “tienen la intención” de realizar una u otra acción. “Tener la intención”, más allá de las sofisticaciones propias de las teorías filosóficas puede tener un referente muy similar en el lenguaje ordinario y en la jerga filosófica. Así los sujetos no tengan una teoría explícita acerca de las acciones intencionales, hay buenos motivos para pensar que cuando alguien, filósofo o sujeto ordinario dice: “Laura tuvo la intención de venir ésta tarde” está refiriéndose a algo muy similar.

Otro ejemplo, donde es quizás más claro, es el de los estudios sobre intuiciones respecto a libre albedrío y determinismo llevados a cabo por Nichols (2004) en niños de 4 años. En varios de estos, Nichols plantea unos escenarios en donde ocurren eventos físicos tales como el hervor de un agua luego de estar un tiempo al fuego y eventos donde hay elecciones morales tales como una niña que roba una dulcería. Después de mostrar éstas viñetas, Nichols (2004) les pregunta a los niños si los sucesos tenían que ocurrir o si pudieron no haber ocurrido. Parece más o menos claro que la expresión “suceso que tenía que ocurrir” mapea satisfactoriamente la expresión “suceso determinado”.

Así, en los casos de trabajo experimental sobre las intuiciones sobre intencionalidad y determinismo no se encuentra el problema que he señalado en el trabajo experimental sobre las intuiciones respecto a la conciencia fenoménica. No hay una palabra del lenguaje ordinario que mapeé

“fenoménico”, haciendo más difícil esta tarea. Ya hemos visto que incluso entre filósofos la definición del término es complicada y muchas veces la manera más fácil de aprehender el concepto es vía ejemplificación. La única palabra que parece servirnos es “sentir” pero sólo parece aplicarse a ciertos estados que los filósofos califican de fenoménicos. Tal es el caso de sensaciones corporales: “sintió frío” (26.400 resultados de esta expresión en una búsqueda en Google.com), o emociones: “sintió tristeza” (8.900 resultados de esta expresión en una búsqueda en Google.com) pero no percepciones visuales “sintió el verde” (sólo 4 resultados de esta expresión en una búsqueda en Google.com) o percepciones auditivas “sintió una canción” (sólo 3 resultados de esta expresión en una búsqueda en Google.com).

Esta intraducibilidad de “fenoménico” a un término del lenguaje ordinario quizás no sea una dificultad de principio para el intento de responder a la pregunta que ha guiado la presentación. Pero sí implica una dificultad adicional en el diseño de los estudios²⁹.

§4.3. Consecuencias para el trabajo teórico

Ahora bien, lo afirmado en esta tesis: la indeterminación de la evidencia en los trabajos experimentales que intentan sistematizar la comprensión ordinaria sobre la conciencia hace que no sepamos si las teorías sobre las que se está discutiendo cumplen el requisito de no artificialidad en el *explanandum*, ¿qué consecuencias tiene para el trabajo teórico sobre la conciencia fenoménica? Dos posiciones generales se podrían adoptar.

Ignorar la cuestión experimental

Los teóricos de los estudios de la conciencia bien pueden ignorar o tachar como poco promisoría esta vía de investigación experimental. Pero, si se aceptan las tesis del programa de la Filosofía Experimental, en particular aquella que señala que apelar a intuiciones implica un compromiso empírico, entonces el teórico en cuestión se enfrentará con un dilema: o bien su investigación no es sobre conciencia fenoménica de la manera tradicionalmente descrita, como señala Dennett (2006) que de hecho sucede, o bien su investigación debe de alguna otra manera sostener la legitimidad del *explanandum* que no sea vía intuiciones, es decir, vía experimentos mentales de concebibilidad,

²⁹ De hecho, Knobe & Prinz (2008) usan el término “sentir” [*feels*] en el diseño de unas pruebas que indagan acerca de la relación atribución de estados conscientes/psicología moral. Según lo mostrado en éste apartado, dado que se usa el término “sentir”, los resultados aplicarán sólo respecto a la atribución de sensaciones corporales y emociones con contenido fenoménico. No para la totalidad de estados que se ha considerado que tienen un contenido fenoménico.

como lo he presentado (§2.3.1.) o vía análisis conceptual donde se involucre el uso cotidiano del concepto. Si la primera es la opción, tal teórico estaría interesado en otro aspecto de la conciencia que, como bien se mostró en una breve enumeración en la Introducción, es interesante pero se escapa de los terrenos de nuestros intereses. Si la segunda es la opción, tal teórico tendría que ofrecer una manera de entender los argumentos de concebibilidad, donde no se haga una apelación a la comprensión ordinaria. Como he señalado (ver §2.3.1.), si no se entienden como haciendo una apelación a la comprensión ordinaria, se podría estar cometiendo una petición de principio.

Continuar el trabajo experimental

Bien se puede continuar con el trabajo experimental de tal manera que se apliquen las restricciones que se señalaron en el apartado anterior. Esto se puede hacer tomando en consideración la sugerencia que se formuló, o bien de alguna otra manera. De cualquier modo, el trabajo experimental podría arrojar en el futuro dos escenarios distintos: que la evidencia muestre que el *explanandum* de las teorías de la conciencia cumplen el RNAE o que no lo cumplen. Si la evidencia muestra que sí se cumple, tendremos una manera de legitimar el *explanandum* de una teoría de la conciencia, podríamos establecerlo como punto común y usarlo como criterio para juzgar el avance de las diversas teorías. El panorama teórico no cambiaría mucho pero se tendría un argumento sólido a la mano a favor de la legitimidad del problema de la conciencia fenoménica (ver §2.3.).

Sin embargo, si la evidencia muestra que no, que el *explanandum* de la conciencia que se ha estudiado no cumple con el requisito de la no artificialidad dado que la comprensión ordinaria de los estados conscientes no corresponde con el concepto de “conciencia fenoménica”, tendríamos algunas consecuencias interesantes.

Una posibilidad que se abriría es que el programa de investigación se haya equivocado al esperar que la explicación teórica, filosófica, y tal vez científica, de la conciencia sea una sola para todo caso de conciencia fenoménica. Es decir, la conciencia fenoménica bien podría *no ser* una clase natural, de tal manera, no tendríamos que esperar una teoría unificada. Así, se abriría la posibilidad de esperar explicaciones diferentes para estados conscientes de tipo distinto. Por ejemplo, una teoría para experiencias perceptivas y una teoría diferente para sensaciones corporales³⁰. La unificación de

³⁰ En este sentido es interesante notar la divergencia entre una teoría de la experiencia perceptiva como la defendida por Humphrey (2006) y una teoría acerca del dolor como la defendida por Grahek (2008). Ambas teorías están científicamente orientadas. Ambas teorías aceptan que los estados conscientes (en un caso ver rojo y en otro caso sentir un dolor) deben entenderse desde una perspectiva que disocie el componentes sensitivo-fenoménico y del

tales estados bajo un sólo rubro se ha establecido a partir del concepto de conciencia fenoménica, pero si la legitimidad de tal concepto queda en entre dicho no habría razón para seguir tomándolos a todos como un paquete homogéneo. Otra posibilidad, es que una vez se considere el *explanandum* de la conciencia fenoménica como ilegítimo, los esfuerzos se concentren en otro tipo de proyecto explicativo de la conciencia, por ejemplo, alguno de los descritos en la Introducción.

cognitivo-perceptivo. Pero difieren en un punto central, en su teoría sobre experiencias perceptivas, Humphrey sostiene que aquello que resulta interesante y valioso para el sujeto es el componente sensitivo-fenoménico, incluso él mismo constituye el valor adaptativo de tales estados conscientes, mientras que en su teoría sobre la naturaleza dual del dolor, Grahek (2008) sostiene, a partir de estudios en pacientes con síndromes disociativos de los dos componentes, que el componente cognitivo-perceptivo es el que realmente importa al sujeto. ¿Son estas dos teorías incompatibles? Sí, si esperamos que debemos tener una teoría que dé cuenta de todos los estados conscientes. Pero no sería un problema en caso de abandonarse tal tarea.

APÉNDICES

Piloto 1 – Literalidad de atribuciones de estados mentales a entidades colectivas

Se realizó una prueba de calificación escala Likert 1 a 5 a una muestra de 32 voluntarios (17 hombres – 15 mujeres) vía web en el servidor de cuestionarios MonkeySurveys.com. El único requisito fue no tener alguna formación en Filosofía, Psicología o Literatura. La edad media de los participantes fue cercana a los 24 años. Se anuló una prueba porque el participante era menor de 18 años. En su computadora los sujetos leyeron el siguiente texto:

“Cuando hablamos a diario decimos cosas que expresan aquello que creemos que sucede en el mundo, pero también decimos cosas que si bien consideramos verdaderas, no pensamos que se correspondan con el mundo.

Así, "Bush fue presidente de Estados Unidos de América" es un enunciado literalmente verdadero, ya que reconocemos que hay una persona de apellido Bush que fue presidente del país Estados Unidos de América. Mientras tanto, "El amor es ciego" lo es sólo de modo figurativo, sólo es un modo de hablar, ya que nadie afirmaría que un sentimiento como el amor podría tener ojos o percepciones.

La siguiente prueba se interesa acerca de cuáles son los límites entre enunciados literales y figurativos.

¡Gracias por participar!”

A continuación encontraban la siguiente instrucción:

A continuación encontrarás una lista de enunciados. Deberás calificar éstas en una escala de 1 a 5. Siendo 1 "Enunciado totalmente figurado" y 5 "Enunciado totalmente literal". Cualquier punto intermedio que consideres más adecuado está permitido.

Y por último la siguiente lista de enunciados, en cada caso la computadora la presentó en un orden aleatorio distinto:

La luna es de queso.

María lloró mares cuando supo que su padre había muerto.

Un triángulo tiene tres lados.

María piensa cambiar de empleo el próximo año.

Pablo decidió ir a cine el sábado en la noche.

Pedro desea ganarse la lotería.

Microsoft piensa implementar un nuevo plan de mercadeo.

Si las cosas van mal, Coca-Cola decidirá vender sus acciones.

Todas las empresas desean mejorar sus ingresos.

Frente a cada enunciado se ubicaron de manera equidistante cinco círculos usados para marcar una calificación en una escala Likert de 5 valores. Siendo (1) “Enunciado totalmente figurado” y (5) “Enunciado totalmente literal”.

Las preguntas de control (1-3) se plantearon para brindar confianza acerca de si la muestra comprendía la diferencia, hecha explícita en el enunciado, entre “enunciado muy literal” y “enunciado muy figurado”, con un resultado de 1,16 en la pregunta 1, y un resultado de 4,68 en la pregunta 3, considero que la muestra entendió con claridad la distinción señalada.

La prueba de significatividad estadística se hizo a partir de una media entre las calificaciones de enunciados cuyos sujetos son entidades individuales (4-6) y las calificaciones de enunciados cuyos sujetos son entidades colectivas (7-9) para cada participante. La prueba, $t(31)=0.29$, $p<.05$, arroja como resultado que no hay una diferencia estadísticamente relevante entre la calificación de literalidad de los enunciados cuyos sujetos son entidades individuales y aquellos cuyos sujetos son entidades grupales.

Piloto 2 - ¿“Ver rojo” se dice de diferentes maneras respecto a diferentes entidades?

Se realizó una prueba de calificación escala Likert 1 a 7 a una muestra de 29 voluntarios (13 hombres – 16 mujeres) vía web en el servidor de cuestionarios MonkeySurveys.com. El único requisito fue no tener alguna formación en Filosofía, Psicología o Literatura. La edad media de los participantes fue cercana a los 23 años. En su computadora los sujetos leyeron el siguiente texto:

(Página 1)

¡Gracias por participar!

En este estudio no hay respuestas correctas o incorrectas, sólo es importante aquello que tú creas. Así que responde estas preguntas sin ayuda de nadie.

Por favor lee con cuidado el escenario que se describirá a continuación, luego responde las tres preguntas y por último dinos tu sexo y edad.

(Página 2)

Jimmy es un robot relativamente simple construido en una universidad estatal. Él tiene una videocámara por ojos, ruedas para moverse y dos brazos con sensores táctiles con los que mueve objetos. Timmy es un joven estudiante normal que hará algunas pruebas para unos científicos. Como parte de un experimento psicológico, Jimmy es puesto junto a Timmy dentro de una habitación vacía excepto por dos cajas azules, dos cajas rojas y dos cajas verdes (las cajas eran idénticas en todos los aspectos excepto en el color).

Los investigadores le transmitieron una instrucción a Jimmy y le pidieron lo mismo a Timmy: “Coloca una caja roja frente a la puerta”. Cada uno se movió, tomó la caja y la puso frente a la puerta sin ninguna dificultad notable.

A continuación deberás responder tres preguntas en una escala de 1 a 7. Siendo 1: Claramente no, 4: No estoy seguro y 7: Claramente sí. También deberás justificar cada respuesta.

¿Jimmy y Timmy ven rojo?

¿Jimmy y Timmy son humanos?

¿Jimmy y Timmy están en la habitación?

El orden de las preguntas fue aleatorio para cada participante. Si bien la primera pregunta fue la importante, la segunda y la tercera fungieron como preguntas de control. Por un lado se quería hacer un control de lectura que mostrara que los participantes reconocían la diferencia entre Jimmy y Timmy. Por otro lado, se quería establecer dos referentes: una pregunta en la que se atribuyera a los dos algo que sólo podría ser atribuido a uno y en la otra en la que se atribuyera a los dos algo que podría ser atribuido a ambos. De este modo, la predicción era que si se defendía (C) (ver apartado §2.2.1.), las respuestas a la primera pregunta tenderían a parecerse a las respuestas a la segunda pregunta. Sin embargo, los resultados muestran que esto no fue el caso. Las medias fueron las siguientes:

¿Jimmy y Timmy ven rojo?	5.41
¿Jimmy y Timmy son humanos?	1.79
¿Jimmy y Timmy están en la habitación?	6.59

Las personas tendieron a responder afirmativamente a la primera pregunta. Así, falla la predicción a la que se comprometería (C).

REFERENCIAS

- Alexander, J. & Weinberg, J. M. (2007). "Analytic epistemology and experimental philosophy", *Philosophy Compass*, 2(1), pp. 56–80.
- Arico, A., Fiala, B., Goldberg, R. & Nichols, S. (por aparecer) "The Folk Psychology of Consciousness".
- Block, N. (1995) "On a confusion about the function of consciousness" en *Behavioral and Brain Sciences*, 18, pp. 227- 247.
- Bloom, P. (2004) *Descartes' Baby: How The Science of Child Development Explains What Make Us Humans*. William Heinemann, Londres.
- Bloom, P. & Veres, C. (1999) "The perceived intentionality of groups" en *Cognition*, No. 71, B1-B9.
- Brendel, E. (2003) "Pompas de intuiciones y el uso adecuado de los experimentos mentales" en *Ideas y Valores*, No. 123, pp. 3 – 23.
- Carruthers, P. (2005) *Consciousness: Essays from a Higher-Order Perspective*, Oxford University Press.
- Chalmers, D. (1995) "Facing Up to the Problem of Consciousness" en *Journal of Consciousness Studies*, 2 (3), pp. 200-219.
- Chalmers, D. (1996) *The Conscious Mind: In Search of a Fundamental Theory*, Oxford University Press.
- Churchland, P. (1985) "Reduction, Qualia, and the Direct Introspection of the Brain States" en *Journal of Philosophy*, Vol. 82, pp. 8 – 28.
- Crick, F. (1994) *The Astonishing Hypothesis: The Scientific Search for the Soul*, Charles Scribner's Sons, New York.
- Cullen, S. (2009) "Survey-Driven Romanticism" en *European Review of Philosophy*, 9, por aparecer.
- Cummins, R. (1998) "Reflection on Reflective Equilibrium" en *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*, DePaul & Ramsey (eds.), Rowman & Littlefield Publishers, Lanham, pp. 113 – 127.
- De Brigard, F. (por aparecer) "Attention, Consciousness and Commonsense", *Journal of Consciousness Studies*.
- Dehaene, S. & Naccache, L. (2001) "Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework" en *Cognition*, Vol. 79, pp. 1 – 37.
- Dennett, D. (1988) "Quining Qualia" en *Consciousness and Contemporary Science*, Marcel & Bisiach (comps.), Oxford University Press, pp. 42 – 77.
- Dennett, D. (1991) *Consciousness Explained*, Little, Brown and Company, Boston.

- Dennett, D. (2001). "Are We Explaining Consciousness Yet?" en *Cognition*, Vol. 79, pp. 221-237.
- Dennett, D. (2005) *Sweet Dreams: Philosophical Obstacles to a Science of Consciousness*, MIT Press.
- Double, R. (1985) "Phenomenal Properties" en *Philosophy and Phenomenological Research*, Vol. 45, No. 3, pp. 383 – 392.
- Fedyk, M. (2009) "Philosophical Intuitions" en *Studia Philosophica Estonica*, 2.2, pp. 54 – 80.
- Feldman, R. (1974) "An Alleged Defect in Gettier Counterexamples", *Australasian Journal of Philosophy*, 52, pp. 68-9
- Foot, P. (1965) "The Problem of Abortion and the Doctrine of Double Effect" en *Oxford Reviews*, 5, No. 8 - 9.
- Gettier, E. (1963) "Is Justified True Believe Knowledge?" en *Analysis*, 23, pp. 121-123.
- Giacino, J. T. (2002) "The minimal conscious state" en *Neurology*, Vol. 58, N. 3, pp. 349-353.
- Goldman, A. (1993) "Consciousness, Folk Psychology and Cognitive Science" en *Consciousness and Cognition*, Vol. 2, pp. 364-382.
- Greenwood, J. (1991) *The Future of Folk Psychology*, Cambridge University Press, Cambridge.
- Hamilton & Sherman (1996) "Perceiving persons and groups" en *Psychological Review*, No. 103, 336-355.
- Hauser, M. (2006) *Moral Minds: How Nature Design Our Universal Sense of Right and Wrong*, Ecco.
- Huebner, B. (2009). "Commonsense concepts of phenomenal consciousness: Does anyone care about functional zombies?" en *Phenomenology and the Cognitive Sciences*, 9 (1), pp. 133-155.
- Humphrey, N. (2006) *Seeing Red: A Study in Consciousness*, The Belknap Press of Harvard University Press.
- Hobbes, T. (2008) *El Leviatán*, Fondo de Cultura Económico, México. [1651]
- Jackson, F. (1982) "Epiphenomenal Qualia" en *Philosophical Quaterly*, Vol. 32, No. 127, pp. 127 – 136.
- Jackson, F. (1998) *From Metaphysics to Ethics: A Defence of Conceptual Analysis*, Oxford University Press.
- Jackson, F. (2003) "Qualia epifenoménicos" en *La Naturaleza de la Experiencia*, Ezcurdia & Hansberg (ed.), UNAM. [1982]
- Kaupinnen, A. (2007) "The rise and fall of experimental philosophy" en *Philosophical Explorations*, 10(2), pp. 95–118.
- Knobe, J. (2003) "Intentional Action and Side Effects in Ordinary Language" en *Analysis*, 63, pp. 190 – 193.
- Knobe, J. & Nichols, S. (2008) "An Experimental Philosophy Manifesto" en *Experimental Philosophy*, Knobe, J. & Nichols, S. (eds.), Oxford University Press.
- Knobe, J. & Prinz, J. (2008) "Intuitions About Consciousness" en *Phenomenology and the Cognitive Sciences*, 7 (1), pp. 67-83.

- Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M. & Damasio, A. (2007) "Damage to the prefrontal cortex increases utilitarian moral judgements" en *Nature*, Vol. 446, pp. 908 - 911.
- Kripke, S. (1980) *The Naming and Necessity*, Harvard University Press.
- Lam, B. (en preparación) "Are Cantonese Speakers Really Descriptivists? Revisiting Cross-cultural Semantics".
- Llinás, R. (2001) *I of the Vortex: From Neurons to Self*, The MIT Press.
- Machery, E., Mallon, R., Nichols, S., & Stich, S. P. (2004). "Semantics, cross-cultural style" en *Cognition*, 92(3), pp. B1–B12.
- Machery, E. (2008) "Understanding the Folk Concept of Intentional Action: Philosophical and Experimental Issues" en *Mind & Language*, 23, pp. 165 – 189.
- Martí, G. (2009) "Against semantic multi-culturalism" en *Analysis*, 69, pp. 42-48
- McCann, H. (2005) "Intentional Action and Intending: Recent Empirical Studies" en *Philosophical Psychology*, 18 (6), pp. 737 – 748.
- Nadelhoffer, T. y Nahmias, E. (2007) "The Past and Future of Experimental Philosophy" en *Philosophical Explorations*, Vol. 10, No. 2. pp. 123 – 149.
- Nagel, T. (1974) "What is it like to be a bat?" en *The Philosophical Review*, 58, 4, pp. 435 – 450.
- Nagel, T. (1996) *Una visión desde ningún lugar*, FCE, México.
- Nahmias, E., Morris, S., Nadelhoffer, T., & Turner, J. (2005) "Surveying freedom: Folk intuitions about free will and moral responsibility", *Philosophical Psychology*, 18(5), pp. 561–584.
- Nichols, S. (2004a) "Folk concepts and intuitions: from philosophy to cognitive science" en *Trends of Cognitive Science*, 8 (11), pp. 514 – 518.
- Nichols, S. (2004b) "The Folk Psychology of Free Will: Fits and Starts", en *Mind and Language*, 19, pp. 473 - 502
- O' Laughlin, M. & Malle, F. (2002) "How people explain actions performed by groups and individuals" en *Journal of Personality and Social Psychology*, No. 82, 33-48.
- Papineau, D. (2002) *Thinking about Consciousness*, Oxford University Press, New York.
- Penrose, R. (1989) *The Emperor's New Mind*, Oxford University Press, New York.
- Ravenscroft, I. (2009) "Is Folk Psychology a Theory?" en *The Routledge Companion to Philosophy of Psychology*, Routledge, New York.
- Searle, J. (1997) *The Mystery of Consciousness*, New York Review of Books.
- Shafir, E. (1998) "Philosophical Intuitions and Cognitive Mechanisms" en *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*, DePaul & Ramsey (eds.), Rowman & Littlefield Publishers, Lanham, pp. 59 – 73.
- Sorensen, R. (1998) *Thought Experiments*, Oxford University Press.
- Sosa, E. (1998) "Minimal Intuition" en *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*, DePaul & Ramsey (eds.), Rowman & Littlefield Publishers, Lanham, pp. 257 – 269.
- Sytsma, J. (por aparecer). "Dennett's Theory of the Folk Theory of Consciousness" en *Journal of Consciousness Studies*.

- Sytsma, J. & Machery, E. (2009) “How to Study Folk Intuitions about Phenomenal Consciousness” en *Philosophical Psychology*, 22(1), pp. 21-35.
- Sytsma, J. & Machery, E. (por aparecer) “Two Conceptions of Subjective Experience” en *Philosophical Studies*.
- Talbot, B. (2009) “Psychology and the Use of Intuitions in Philosophy” en *Studia Philosophica Estonica*, 2.2, pp. 157 – 176.
- Tye, M. (1996) *Ten Problems of Consciousness: A Representational Theory of Phenomenal Mind*, The MIT Press.
- Unger, P. (1968). “An Analysis of Factual Knowledge”, *Journal of Philosophy*, 65, pp. 157-70.
- Uttich, K. & Lombrozo, T. (en prensa) “Norms Inform Mental States Ascriptions: A Rational Explanation for the Side-Effect Effect”, *Cognition*.
- Williamson, T. (2008) *The Philosophy of Philosophy*, Wiley-Blackwell.
- Wisniewsky, E. (1998) “The Psychology of Intuitions” en *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*, DePaul & Ramsey (eds.), Rowman & Littlefield Publishers, Lanham, pp. 45 – 57.
- Zeman, A. (2003) “What is consciousness and what does it mean for the persistent vegetative state?” en *Advances in Clinical Neuroscience and Rehabilitation*, Vol. 3, N. 3.