



UNIVERSIDAD NACIONAL DEL SUR

TESIS DE DOCTORADO EN FILOSOFÍA

Conciencia, Cualidades y Fenómenos:
La Superación de la Impotencia Reduccionista

Jorge Gabriel Mux

BAHIA BLANCA

ARGENTINA

2014

PREFACIO

Esta tesis se presenta como parte de los requisitos para optar al grado académico de Doctor en Filosofía, de la Universidad Nacional del Sur y no ha sido presentada previamente para la obtención de otro título en esta Universidad u otra. La misma contiene los resultados obtenidos en investigaciones llevadas a cabo en el ámbito de Departamento de Humanidades de la Universidad Nacional del Sur durante el período comprendido entre el año 2006 y el año 2014, bajo la dirección de los Dres. Oscar Esquisabel y Gustavo Bodanza.

.....



UNIVERSIDAD NACIONAL DEL SUR
Secretaría General de Posgrado y Educación Continua

La presente tesis ha sido aprobada el/...../....., mereciendo la calificación de (.....).

Resumen

El presente trabajo se propone como un análisis de las características fenomenológicas de la conciencia, y un examen de las razones del continuo fracaso por explicarlas reductivamente. Una explicación exitosa debería mostrarnos una continuidad entre el reino psíquico y el reino físico, y debería servirnos para unificar conceptualmente la teoría que trata ambos dominios. Sin embargo, el fracaso en esta unificación nos deja con una disunidad explicativa: por un lado, tenemos explicaciones puramente fenoménicas y por otro, explicaciones funcionales, físicas o mecánicas. La imposibilidad de reducir ambos dominios atenta contra el ideal de unidad de las explicaciones científicas, un ideal que parece estar detrás de toda búsqueda de reducción.

El presente trabajo se divide en dos partes. En la **primera parte** haremos una evaluación crítica de las principales respuestas del filósofo John Locke acerca del problema de la unidad de la explicación mecanicista del mundo y la compatibilidad de dicha explicación con los estados de conciencia. El mecanicismo clásico intenta explicar el mundo mediante la reducción de todos los hechos del mundo a pautas mecánicas, pero se enfrenta a un problema insoluble cuando intenta operar esta reducción a los hechos mentales, o cuando intenta conectar el ámbito mental con el ámbito físico. Para conectar ambos ámbitos, la propia noción de 'idea' debe cabalgar ambiguamente entre dos mundos: es una entidad mental, interna, que sin embargo contiene una huella causal externa, lo cual por principio contraría la propia noción de 'interno'. Mostraremos que, a pesar de los esfuerzos de la modernidad, el programa mecanicista clásico no puede encontrar una continuidad entre el aspecto fenoménico de la mente y el aspecto mecanicista de los cuerpos, pues aun cuando propone que en la idea hay una huella causal, tal huella causal no deja de ser inevitablemente fenoménica.

El fracaso de la reducción del aspecto interno al aspecto externo no sólo vale para el siglo XVII, sino también para el siglo XXI. En la **segunda parte** analizaremos algunos de los intentos de explicación de los fenómenos mentales en términos funcionalistas y/o eliminativistas. El funcionalismo asume que una propiedad –como el color, por ejemplo- queda totalmente explicada si se especifican los mecanismos funcionales

operantes subyacentes en un proceso. El eliminativismo afirma que la propia fenomenología resulta errónea y ociosa, y en rigor no representa un problema genuino, y por lo tanto debe eliminarse. Veremos que estos intentos llegan a ciertos callejones sin salida, pues los esquemas funcionales que se presentan como explicativos y reductivos de la naturaleza cualitativa de una percepción, contienen ellos mismos importantes elementos fenoménicos, con lo cual la explicación –tal como ocurría con el mecanicismo moderno- no puede establecer una continuidad entre el aspecto fenoménico y el mundo externo.

Propondremos que el principal problema de las explicaciones reductivas acerca de la conciencia es que se valen de una fuerte demarcación entre el aspecto interno (mental) y el aspecto externo (físico, funcional) y que una mejor explicación de los hechos conscientes puede darse a partir de una *teoría enactiva*, según la cual la conciencia no es un hecho “interior”, sino un hecho ineliminablemente ecológico.

Abstract:

This work proposes an analysis of the phenomenological features of consciousness, and a review of the reasons for the continued failure by reductive explanations. A successful explanation should show continuity between the psychic realm and the physical realm, and should serve to conceptually unify the theory that treats both domains. However, this unification failure leaves us with an explanatory disunity: first, we have purely phenomenal explanations and secondly, functional, physical or mechanical explanations. The inability to reduce both domains attempts against the ideal of unity of scientific explanations, an ideal that seems to be behind any reduction. This work is divided into two parts. In the first part we will make a critical assessment of the main responses of the philosopher John Locke about the problem of the unity of the mechanistic explanation of the world and the compatibility of the explanation of states of consciousness. The classic mechanicism attempts to explain the world by reducing all facts of the mechanical world patterns, but faces an intractable problem when it tries to operate this reduction to mental events, or when tries to connect the mental realm to the physical realm. To connect the two areas, the very notion of 'idea' must ride

ambiguously between two worlds: is a mental, internal entity, which however contains an external causal footprint, which in principle contradicts the notion of 'internal'. We show that, despite the efforts of modernity, the classical mechanistic program cannot find a continuity between the phenomenal aspect of mind and the mechanistic aspect of the bodies, because even when he proposes the idea that there is a causal footprint, such causal footprint is no longer inevitably phenomenal. The failure of the reduction of the internal to the external appearance looks not only applies to the seventeenth century, but also for the XXI century. In the second part we will analyze some of the attempts to explain mental phenomena in functionalist and / or eliminativists terms. Functionalism assumes a property-like color, for example, is fully explained if there is operative functional mechanisms underlying a specified process. Eliminativism claims that phenomenology itself is wrong and idle, and in fact does not represent a genuine problem, and therefore should be removed. We will see that these attempts come to some dead ends, because the functional diagrams presented as explanatory and reductive of the qualitative nature of perception, contains important phenomenal elements themselves, which explanation-such as with the modern mechanism - can not establish a continuity between phenomenal character and the external world. We propose that the main issue of reductive explanations of consciousness is the use of a sharp demarcation between the internal (mental) and external appearance aspect (physical, functional) and a better explanation of conscious events can occur from an enactive theory, according to which consciousness is not an "internal" fact but an ecological one.

A Isabella, Irma, Jorge Ricardo y Alicia.

Agradecimientos.

Deseo manifestar, en primer lugar, mi profundo agradecimiento a los Dres. Oscar Esquisabel y Gustavo Bodanza, quienes guiaron y supervisaron el proceso de redacción de este trabajo.

Agradezco también al Dr. Marcelo Auday, cuyos desinteresados consejos y sugerencias ayudaron a delinear y decidir aspectos cruciales de la temática elegida, además de facilitarme bibliografía clave para desarrollar las investigaciones.

Destaco y agradezco la amigable presencia tutelar de la Dra. Zulma Mateos, quien durante mucho tiempo incentivó y apoyó activamente el camino que concluye en este trabajo, aportándome sugerencias inestimablemente valiosas, además de haberme permitido trabajar con ella. Es muy posible que, sin la formación recibida por ella, este trabajo nunca hubiera sido realidad. Agradezco también a los Dres. Juan Manuel Torres y Jorge Roetti, de quienes aprendí muchísimo.

No me olvido de mis colegas, amigos y compañeros de trabajo: Hipólito Hasrun, María Inés Silenzi, Rebeca Canclini, José Schmidt y Ayelén Sánchez.

Por último, quiero agradecer a mi abuela Lita quien, con sus noventa años, nunca deja de alentarme para la conclusión de este trabajo.

Conciencia, Cualidades y Fenómenos: La superación de la impotencia reductivista.

Jorge Mux

Contenido

Conciencia, Cualidades y Fenómenos: La superación de la impotencia reductivista.	1
Introducción	5
Primera parte: Cualidades, ideas y la reducción mecanicista de la modernidad.	11
1.1 La <i>disunidad</i> de la idea.	15
1.2. Cualidades.	26
1.2.1. Propiedades generales de la experiencia: tres requisitos que deben satisfacer las cualidades primarias.	34
1.2.2. Tendiendo el puente.	47
1.2.3. Aberraciones y más aberraciones.	53
1.3. Simetría de cualidades, asimetría de ideas y reducción.	60
1.4. El “micrango” y la relatividad de la reducción entre cualidades.	64
1.5 La semejanza entre cualidades primarias e ideas.	73
1.5.1. La tesis corpuscular	81
1.5.2. La tesis de la convergencia geométrica.....	83
1.5.3. ¿Cómo entender la semejanza (<i>resemblance</i>)?	84
1.6. La desemejanza de las cualidades secundarias y las ideas: apelación a la arbitrariedad divina.	86
1.7. ¿Toda cualidad es secundaria?	92
1.8. Frege y su lucha contra el representacionalismo.....	101
1.9. Balance: el realismo por representación versus el realismo intencionalista.....	120
Segunda parte: Pareceres, qualia y la reducción de los estados cualitativos durante fines del siglo XX y principios del XXI.	125
2.1. Acerca del parecer.....	135
2.2. ¿No le parece que carece de parecer?.....	140
2.3 Hablando acerca del inexistente anillo rosado del que no tiene sentido hablar.....	147
2.4. “Dennettizar” las explicaciones o “Nagelizarlas”	168
2.5. Los <i>qualia</i> no son <i>lo que parece</i> . (O: ¿le parece a usted que a mí me parece que el césped es verde?).....	176
2.6. Dennettizando la explicación sobre los qualia: la explicación de Jeff Foss.....	184
2.7. ¿Está Maxwell sistemáticamente equivocado en su percepción de colores?	193
2.8. ¿Sueñan las palomas con direcciones aéreas ultravioletas?	207
2.9. Acerca del parecer: ¿Adónde hemos llegado?.....	221

2.10 Modelos explicativos disunificados.....	232
Conclusiones: ¿Dónde está la conciencia? Un diagnóstico especulativo	251
Cabalgando por encima de las analogías fallidas: <i>Enactividad</i> de la conciencia o <i>exoconciencia</i>	259
Referencias.....	269

Introducción

Una buena parte de los problemas en filosofía de la mente estriba en la dificultad para elaborar una base explicativa de la conciencia que sea, a la vez, científicamente aceptable, explicativamente sustanciosa, argumentativamente convincente y epistémicamente satisfactoria: queremos que una explicación de la conciencia nos *llene* y no nos deje con (la sensación de que queda) algo sin explicar. Como ha ocurrido en la filosofía de la mente durante buena parte del siglo XX, las explicaciones acerca de fenómenos mentales han transcurrido dejando de lado aspectos cruciales de la naturaleza de lo mental. Uno de estos aspectos cruciales (y quizás, a nuestro juicio, el más decisivo y característico) es el aspecto fenomenológico. Las propias características fenomenológicas –lo que puede considerarse la *marca de identidad* de lo mental- han sido en muchos casos tratadas como un hecho menor y marginal que merece desaparecer en una explicación reductiva exitosa. Las explicaciones que reducen la fenomenología a acaecimientos de bajo nivel dan la impresión de *no estar explicando* lo que debía ser explicado: si se utiliza una estrategia explicativa mediante la cual el *explanandum* fenoménico desaparece, transformado en *explanantes* de naturaleza puramente funcional, se tiene la impresión de que se ha realizado una especie de pase mágico conceptual: “*Pides que te expliquen A. Transforman A en B, C, D. Luego explican B, C, D y dicen que te han explicado A*”. Si se sospecha que ciertas explicaciones reductivas nos venden agua por vino, tendremos una insatisfacción epistémica que no podrá saciarse con la búsqueda de *mejores* reducciones. Pero, ¿es la insatisfacción epistémica un *buen* argumento antirreduccionista? ¿Debe la ciencia lidiar con las insatisfacciones psicológicas? Hay buenos motivos para sospechar que la insatisfacción epistémica derivada de las explicaciones reductivas en los problemas de la conciencia no es puramente un capricho de filósofos con mentes retorcidas, sino una genuina demanda que requiere de la seria atención tanto científica como filosófica. La conciencia ha generado problemas que parecen desconcertantes en el mejor de los casos, y sorprendentemente intratables en el peor. Cuando algunos filósofos proponen una solución (o una disolución) de ciertos problemas, aparecen de

inmediato otras voces interponiendo objeciones serias y sustanciales cuestionando no sólo la solución propuesta, sino la visión de fondo sobre la cual se elaboran las soluciones. Las explicaciones tambalean entre el optimismo científico –los estados conscientes pueden explicarse *sin resto* mediante leyes *actualmente disponibles*, y mediante mecanismos *empíricamente accesibles*- y el pesimismo más radical –ninguna explicación es posible, ni siquiera imaginable.

El presente trabajo se propone como un análisis de las características fenomenológicas de la conciencia, y un examen de las razones del continuo fracaso por explicarlas reductivamente. Estas características de la conciencia han recibido numerosos (y no necesariamente coextensivos) nombres a lo largo de la historia del pensamiento: Ideas, *qualia*, vivencias, notares, pareceres, cualidades secundarias. Las características fenomenológicas son los elementos de la experiencia consciente. Usted está leyendo este texto, y tiene una particular impresión del fondo blanco de la página. Ese color blanco que impacta en sus ojos tiene una peculiaridad indefinible pero inconfundible. Su experiencia de “ver algo blanco” no puede ser comunicada a una persona ciega, y usted tendría la sensación de que, si intentara comunicarla, quedaría afuera de esa explicación precisamente el hecho que deseaba explicar: la característica *intrínseca* de la experiencia de estar viendo blanco. Ese blanco que está en su percepción (en su mente) y que usted no puede comunicarle a un ciego (un ciego que jamás haya visto ese color) es una característica fenomenológica de su conciencia. Muchas de las explicaciones acerca de los estados cualitativos de la conciencia han sido vistas como la explicación que daría una persona que es *ciega* a esos estados cualitativos: podría explicar causal o funcionalmente el proceso de *ver blanco*, pero dejando afuera, precisamente, el hecho de mismo de la blancura “en la conciencia”.

Cuando hablamos de reducir las características fenoménicas a características funcionales, físicas o mecánicas, buscamos una explicación según la cual el acaecimiento fenoménico Ψ es del mismo tipo que cierto conjunto de acaecimientos ϕ (estos últimos acaecimientos serán funcionales, físicos o mecánicos, dependiendo del tipo de entidades y relaciones que acepte la teoría reductiva). A su vez, dado que las características fenoménicas Ψ son internas (forman el contenido de una conciencia), esperamos que una reducción exitosa nos muestre cómo esas características se

conectan con (o son parte de) el mundo externo φ . Una explicación exitosa debería mostrarnos una continuidad entre Ψ y φ , y debería servirnos para unificar conceptualmente la teoría que trata ambos dominios. Sin embargo, el fracaso en esta unificación nos deja con una disunidad explicativa: por un lado, tenemos explicaciones puramente fenoménicas y por otro, explicaciones funcionales, físicas o mecánicas. La imposibilidad de reducir ambos dominios atenta contra el ideal de unidad de las explicaciones científicas, un ideal que parece estar detrás de toda búsqueda de reducción.

Los estados fenoménicos se presentan como intratables. Si se los trata, el tratamiento es insatisfactorio. Si es satisfactorio, es insuficiente.

El presente trabajo se divide en dos partes. En la primera parte haremos una evaluación crítica de las principales respuestas del filósofo John Locke acerca del problema de la unidad de la explicación mecanicista del mundo y la compatibilidad de dicha explicación con los estados de conciencia. El mecanicismo clásico es un excelente ejemplo de un intento de reducción entre estados fenoménicos y relaciones mecánicas entre corpúsculos. El mecanicismo clásico intenta explicar el mundo mediante la reducción de todos los hechos del mundo a pautas mecánicas, pero se enfrenta a un problema insoluble cuando intenta operar esta reducción a los hechos mentales, y, consecuentemente, cuando intenta conectar causalmente las características de lo mental con las características mecánicas de la realidad externa. La filosofía de John Locke está enmarcada en la visión cartesiana del mundo, para el cual hay dos tipos de sustancia: la sustancia mental y la sustancia extensa. La mente contiene ideas. Existe una brecha entre ambos tipos de sustancia, y la intención de Descartes y Locke es encontrar una explicación de la relación entre ambas sustancias que pueda enmarcarse en el proyecto mecanicista del siglo XVII. Las herramientas conceptuales para llevar a cabo esta explicación son las nociones de *cualidades primarias* y *cualidades secundarias*. Las cualidades son intermediarios entre los cuerpos materiales y las ideas en la mente. John Locke encarnará de manera trágica el péndulo entre el optimismo científico y el pesimismo radical: por un lado, confía en los avances de la ciencia para hacer cada vez más pequeña esta brecha. Por otro lado, es escéptico con respecto a que la ciencia pueda algún día conocer la naturaleza última de la realidad.

Por un lado, pretende dar una explicación de los estados cualitativos de la conciencia mediante las nociones de *cualidades primarias y secundarias*. Por otro lado, se muestra vacilante con respecto a la diferenciación entre tipos de cualidades, y entre cualidades e ideas. Por un lado confía en que la percepción se guía mediante leyes. Por otro lado, supone que las explicaciones acerca de los estados perceptivos son producto de la arbitrariedad divina, lo cual excluye a todo tipo de explicación nómica. Veremos que los fundamentos del programa filosófico más ambicioso del siglo XVII no pueden explicar de manera satisfactoria la relación entre la mente y el mundo. Para poder reducir el aspecto fenoménico (lo que Descartes y Locke denominan 'ideas') a pautas puramente mecánicas, es necesario mostrar alguna conexión causal entre el fenómeno mental y los cuerpos. Esta noción es la base del *realismo por representación*: las ideas, estados fenoménicos puramente mentales, tienen un vestigio de la realidad externa que les dio origen. Pero para lograr conectar el ámbito externo y el interno, la propia noción de 'idea' debe cabalgar ambiguamente entre dos mundos: es una entidad mental, interna, que sin embargo contiene una huella causal externa, lo cual por principio contraría la propia noción de 'interno'. Mostraremos que, a pesar de los esfuerzos de Descartes y Locke, el programa mecanicista moderno no puede encontrar una continuidad entre el aspecto fenoménico de la mente y el aspecto mecanicista de los cuerpos externos, pues aun cuando propone que en la idea hay una huella causal, tal huella causal no deja de ser inevitablemente fenoménica. Esta discontinuidad es la base del fracaso del intento de reducción del aspecto fenoménico al aspecto mecánico.

Como cierre de esta primera sección –y a modo de apéndice– mostramos el enfoque de Gottlob Frege, quien intenta diferenciarse del realismo por representación separando tajantemente el aspecto puramente representativo (*Vorstellung*, la idea como fenómeno mental) del sentido (*Sinn*, el aspecto objetivo de la idea), una posición que se presenta como alternativa y que, a despecho de sus propios problemas epistemológicos, se muestra como una salida promisorio al problema de la reducción.

El fracaso de la reducción del aspecto interno al aspecto externo no sólo vale para el siglo XVII, sino también para el siglo XXI. En la segunda parte analizaremos algunos de los intentos de explicación de los fenómenos mentales en términos funcionalistas y/o eliminativistas. El funcionalismo asume que una propiedad –como el color, por

ejemplo- queda totalmente explicada si se especifican los mecanismos funcionales operantes subyacentes en un proceso. El eliminativismo afirma que la propia fenomenología resulta errónea y ociosa, y en rigor no representa un problema genuino, y por lo tanto debe eliminarse. Veremos que estos intentos llegan a ciertos callejones sin salida, pues los esquemas funcionales que se presentan como explicativos y reductivos de la naturaleza cualitativa de una percepción, contienen ellos mismos importantes elementos fenoménicos, con lo cual la explicación –tal como ocurría con el mecanicismo moderno- no puede establecer una continuidad entre el aspecto fenoménico y el mundo externo. Si el *explanandum* fenoménico es reducido mediante un *explanans* igualmente fenoménico, la reducción no se ha logrado. La impotencia explicativa de Locke es un ejemplo que se repite en otros contextos, en otras épocas y con andamiajes conceptuales diferentes. Existen, sin embargo, ambigüedades semánticas que se mantienen casi intactas a pesar de los cambios conceptuales; existen objetivos que no se han modificado a pesar del paso de los siglos. La estrategia reductiva sigue mostrando su impotencia, y los estados fenoménicos que pretendíamos hacer desaparecer mediante una explicación eliminativista o funcionalista, reaparecen a cada segundo delante de nuestras narices - y nuestros sentidos en general.

Pero, ¿por qué los aspectos fenoménicos se resisten a ser reducidos? ¿Quizás porque no se ha encontrado aun un marco conceptual reductivo adecuado? ¿Tal vez porque la propia fenomenología es irreductible? ¿O porque la reducción es una mala estrategia explicativa? Se argumenta que, a despecho de lo que se ha creído durante algunas décadas –que la conciencia es un caso *especial* de irreductibilidad-, el aspecto fenoménico no es el único ámbito irreductible: las discontinuidades explicativas están en todas partes. La conciencia es un caso más entre un sinnúmero de hechos que se resisten a ser modelados en una explicación reductiva unificada y sin resto explicativo. Desde este punto de vista, en la explicación de un hecho es posible optar, entonces, por una multitud de modelos no ulteriormente unificables, y por lo tanto la explicación de la propia conciencia debería admitir más de un modelo.

Propondremos que el principal problema de las explicaciones reductivas acerca de la conciencia es que se valen de una fuerte demarcación entre el aspecto interno

(mental) y el aspecto externo (físico, funcional) y que una mejor explicación de los hechos conscientes puede darse a partir de una *teoría enactiva*, según la cual la conciencia no es un hecho “interior”, sino un hecho ineliminablemente ecológico. Si se entiende a la conciencia como algo que ocurre “en el interior” de ciertos seres (humanos y animales), entonces es posible que se esté malinterpretando el *explanandum*. Se ha interpretado la expresión “ser contenido de conciencia” como algo que ocurre en una recóndita interioridad. Pero los contenidos de conciencia no tienen por qué ser, ellos mismos, algo *interno*. Quizás la conciencia sea algo externo y esté difundida en todos los hechos de los que somos conscientes. Aceptada esta propuesta, no necesitamos aplicar una estrategia reductiva para explicar los fenómenos: fenómeno interno y realidad exterior coinciden, en tanto que los hechos de mi conciencia son, literalmente, los hechos del mundo. Quizás la imposibilidad de reducción que encontramos en el mecanicismo clásico, como en el funcionalismo contemporáneo, se debe a que ambos modelos necesitan de una demarcación fuerte entre estado fenoménico (interno) y estado mecánico / funcional (externo). Desde tales perspectivas, la conciencia siempre está “del lado interno”, separada de la externidad a la que se pretende reducir. En cambio, si entendemos a la conciencia como algo que no puede conceptualizarse de acuerdo a la dicotomía ‘externo – interno’, quizás sea posible pensarla como un hecho difundido en el mundo –mundo del cual la propia conciencia forma parte. Si estas especulaciones son aceptables, se podrá entender que los intentos de reducción han buscado a la conciencia en un lugar inadecuado: han situado a los estados fenoménicos en una sede unificada y unificadora cuyos elementos deberían encontrarse en algún *locus* mecánico – funcional, y por ello pretenden establecer los puentes reductivos entre ellos, a fin de hallar una explicación de la naturaleza de los fenómenos conscientes que pueda unificarse con las explicaciones mecánico - funcionales. Pero si la conciencia se difunde a través de todo aquello que es su objeto no tiene sentido seguir buscándola exclusivamente en algún recóndito lugar de nosotros mismos ni reducirla a un conjunto de relaciones mecánicas o funcionales.

Primera parte: Cualidades, ideas y la reducción mecanicista de la modernidad.

Una buena parte de los problemas filosóficos relacionados con la naturaleza de la conciencia nacen en la modernidad, con el pensamiento de René Descartes (1596 – 1650). Una característica saliente y notoria del proyecto cartesiano es la aceptación del dualismo sustancial: lo mental y lo físico son dos reinos ontológicos diferentes (sustancia pensante y sustancia extensa), cuya ontología se conoce debido al *modo de acceso* peculiar de cada uno. La sustancia pensante se caracteriza por poseer contenidos de conciencia, y cualquier contenido de conciencia es una *idea*. La sustancia extensa se comporta de modo mecánico, y puede ser íntegramente explicada mediante un estudio exhaustivo de la acción por contacto de los cuerpos (inercia, presión y choque). Definidos los espacios ontológicos de cada una de las sustancias, es necesario esclarecer el estatus de las relaciones entre ellos. Para ello tanto Descartes como Galileo emplearán una clasificación que ya se puede encontrar en Demócrito y cuya formulación más cuidada aparece en John Locke: la distinción entre cualidades primarias de los cuerpos y cualidades secundarias. Las cualidades primarias están propiamente en los cuerpos, aunque las cualidades secundarias sólo son poderes en las cosas que, en contacto con determinados sentidos de los sujetos percipientes, generan un cierto tipo de estado mental que no tiene relación con los cuerpos externos. Uno de los más grandes proyectos del programa mecanicista del siglo XVII estuvo dado por el intento de reducir las cualidades secundarias a cualidades primarias. Lograda esta reducción, se puede establecer la continuidad entre las ideas y los cuerpos, y se abre la puerta para construir una explicación unificada de lo mental y de lo físico. John Locke (1632 – 1704) encarnará de manera dramática el análisis minucioso de las nociones de cualidad primaria y cualidad secundaria; la confianza en la filosofía corpuscular - mecanicista y la impotencia de una reducción entre cualidades e ideas. Aunque Descartes proporciona el marco conceptual desde donde se analiza la relación mente – cuerpo, Locke examinará los límites de estas estructuras conceptuales y tratará de encontrar, infructuosamente, un canal de conexión entre las ideas y el mundo. Su esfuerzo por caracterizar la naturaleza de las cualidades primarias y secundarias (y de estas últimas en términos de las primeras) lo convierte en una figura insoslayable para tratar el problema de la conciencia, y las limitaciones con las que se ha enfrentado su programa reaparecen en intentos contemporáneos de explicaciones reductivas.

En esta primera parte analizaremos algunos de los problemas semánticos, metodológicos y ontológicos de la noción de “idea” cartesiana y los resultados filosóficos de Locke con respecto a la conexión entre conciencia y mundo. El fracaso de Locke al respecto es un síntoma de otros fracasos explicativos del siglo XX y XXI, como veremos en la segunda parte. Al final de esta primera parte contrastaremos los resultados del *realismo por representación* de Locke con la postura de Gottlob Frege, para quien la conexión entre lo mental y el mundo externo no debe darse en términos de representación ni de contenido causal.

1.1 La *disunidad* de la idea.

‘Idea’ es una de esas palabras cuya polisemia a veces inadvertida pone en riesgo cualquier intento de utilización provechosa. A lo largo de la historia, la palabra ‘idea’ ha significado tantas cosas, tanto en el ámbito de la filosofía y de la ciencia como del pensamiento ordinario, que si se evita una dilucidación cuidadosa de sus connotaciones semánticas es posible que se malogre la tarea argumentativa.

Es difícil establecer una referencia unívoca para la palabra ‘idea’, especialmente si se tiene en cuenta que el propio estatuto ontológico de cada una de las referencias posibles ha sido y sigue siendo discutido. Sin embargo, se asume tradicionalmente que el término *idea* puede referirse a algo de naturaleza *lógica, ontológica, trascendental o psicológica*, sin que estos cuatro aspectos sean excluyentes y claramente diferenciables entre sí (Cepedal, 1976, págs. 213 - 216). A lo largo de la historia parece predominar alguno de estos aspectos, en desmedro de los otros. Nuestra intención es mostrar que la palabra ‘idea’ no es simplemente un término polisémico: la propia referencia de lo que se entiende por ‘idea’ es un hecho no unívoco: de allí que no sólo la palabra ‘idea’ sea polisémica, sino que “las ideas” como entidades son heterogéneas en lo que respecta a su ontología.

Tomaremos la palabra ‘idea’ tal como fue entendida en el marco de la modernidad cartesiana. De los cuatro aspectos mencionados arriba, el empirismo inglés –de la mano de Descartes- entenderá a la idea en su sentido psicológico: como representación en una mente y como el elemento básico de toda actividad mental. Descartes toma la palabra idea en dos sentidos: uno general y otro restringido. En su sentido general, idea es “todo lo que el espíritu concibe de modo inmediato” (1957, pág. 147). En su sentido restringido, las ideas “son como imágenes de las cosas, y a éstos solos conviene con propiedad el nombre de idea” (1957, pág. 33). Locke definirá a la idea como “el objeto del entendimiento cuando un hombre piensa” ([1710] 1956, pág. 21). Para David Hume, las ideas son “representaciones exactas de las impresiones que he sentido” (1981, pág. 89) (refiriéndose a las ideas *simples*). La noción de la concepción *inmediata* de una idea en la definición cartesiana debe contraponerse a otras clases de conocimiento puramente *inferencial*. Para realizar una inferencia, es

necesario partir de una idea (inmediata) para llegar a otra idea (mediata). Esta segunda idea -si bien es mediata con respecto a la idea que sirvió de punto de partida- una vez inferida puede ser conocida inmediatamente, de modo intuitivo. Sólo el *acto de relacionar* ideas es no-inmediato. La intuición es garantía de infalibilidad; el error sólo puede darse en la inferencia. Y dado que el lenguaje (en su función primaria, puramente privada) consiste en asignar una marca a las ideas –y, por lo tanto, hacer un trabajo de relación entre ellas-, podemos cometer errores en el acto de marcar la idea con un signo lingüístico¹.

Tomemos el sentido restringido de la definición cartesiana de idea: una idea es como una imagen de la cosa. Una idea representa a una cosa del mismo modo en que un cuadro representa un paisaje: ese parece ser el sentido más natural e inmediato de la palabra ‘imagen’: la imagen pictóricamente isomorfa. En principio, podemos vernos inducidos a esta interpretación por las siguientes palabras de Descartes: “En verdad, el uso que debemos hacer de la intuición intelectual lo conocemos ya al compararlo con la visión ocular” (Regla IX)² Sin embargo, Descartes ha utilizado este término en un sentido menos restringido que el del isomorfismo pictórico: si las ideas sólo fuesen como fotos o como cuadros, hay una infinidad de cuestiones que jamás podrían ser objeto de pensamiento. La imagen a la que hace referencia Descartes debe ser escudriñada con la “luz natural” de la mente (Los Principios de la Filosofía, 1951, principios XXVIII, XXX, XLIV; Meditaciones metafísicas: segunda meditación): la luz natural proporciona la capacidad de inspeccionar el espíritu. Dicha inspección es un tipo de visión mental que difiere de la visión perceptiva, en tanto sus objetos no han de ser sólo imágenes sensoriales, sino también objetos para las cuales no existe la posibilidad de formar tales imágenes: mientras el ojo percibe un grupo abigarrado de colores verdes y marrones, la luz natural del intelecto percibe diez árboles: nada en la percepción ocular suministra el número diez; no hay manera de representarse pictóricamente al diez, ni a las operaciones matemáticas asociadas. Sin embargo, debemos pensar que existe un proceso análogo al de la percepción visual aplicable a

¹ En su función secundaria, el lenguaje es un medio público de comunicación de ideas. Como la función secundaria es subsidiaria de la función primaria, es patente que la función pública cabe mayor posibilidad de error que en la privada.

² En Descartes (1967), pág. 71.

las ideas que no poseen contenido pictórico. Ian Hacking, al respecto de esta “visión sin representación pictórica” siente que

[n]os resulta muy difícil comprender este desaparecido concepto de visión mental. No se pensaba que las ideas eran como imágenes. Por el contrario, el razonamiento más elevado tiene objetos, tales como Dios o la voluntad, de los cuales no podemos, en principio, formarnos ninguna imagen. (Hacking, 1979, pág. 46)

La mente posee una *accesibilidad total* con respecto a sus ideas, de modo que, dada una idea presente al espíritu, este conoce de modo inmediato que dicha idea está presente. La total accesibilidad de las ideas, en tanto son vivenciadas, debe oponerse a la opacidad de los *acaecimientos*. Mientras las vivencias poseen características puramente subjetivas, los acaecimientos pueden definirse como “situaciones o condiciones objetivas, con características objetivas, con una ubicación determinada en el espacio y el tiempo” (García Carpintero, 1996, pág. 61). Las propiedades de lo objetivo son: la *intersubjetividad*: un mismo acontecimiento es accesible a dos o más individuos; la *sustantividad*: los acaecimientos pueden darse sin que nadie se los represente; *fisicidad*: los acaecimientos pueden describirse con precisión en términos científicos, y *normatividad*: el acaecimiento suministra las reglas para decidir cuándo la percepción es exitosa (Hacking, 1979, págs. 61-2). Pero, dado que sólo tenemos contacto con nuestras propias ideas –y no, al menos de modo inmediato, con el mundo exterior-, la afirmación de que hay acaecimientos independientes de las ideas es la conclusión de un argumento y no una afirmación sustentada por la vivencia de una idea. Por el contrario, nuestras ideas –de conocimiento inmediato- al ser *vivencias* poseen características opuestas a las del *acaecimiento*. La característica principal de la vivencia es que, en vez de *acontecer* (esto es, tener una existencia sustancial e intersubjetiva), su modo de existencia es un *notar*³: una idea sólo puede existir en tanto es *notada* por un alma. En otras palabras, una idea sólo existe como objeto de contemplación de una mente. Las características específicas de una vivencia son:

³“(…) Para exponer con claridad una teoría del contenido proposicional de los estados mentales (...) necesitamos un nuevo término técnico que, sin ambigüedad, haga referencia a un estado mental que se dirija a vivencias y no a estados de cosas. Utilizaré ‘notar’ para este fin, y hablaré de notares para referirme a actos o estados mentales en que un individuo nota las características de una de sus vivencias”. García Carpintero, *op. cit.*, pág. 67.

1) Privacidad: Dos personas pueden, en principio, tener dos vivencias cualitativamente similares, pero no podrían tener ambas una vivencia numéricamente idéntica.

2) Transparencia: no hay vivencias que se den sin ser notadas. De hecho, todo el ser de las vivencias consiste en su *notar*.

3) Irreducibilidad: una vivencia no puede ser “mejor conocida” por un análisis o descripción en términos precisos: la única manera de conocer una vivencia es *notando* esa vivencia. Cualquier descripción por exhaustiva que fuese no puede dar un conocimiento de la específica cualidad de esa vivencia, la cual sólo puede darse a partir del *notar*.

4) Incorregibilidad: no hay notares falsos o frustrados. (García Carpintero, 1996, pág. 73)

Cuando la luz natural, mediante una inspección del espíritu, encuentra sus objetos (las ideas), les otorga el estatus de “vivencias”, esto es, las ideas se hacen presentes al alma. Si el individuo posee una idea, tal idea puede ser notada (a esto llamábamos *accesibilidad total*: Descartes encuentra absurdo que alguien pueda poseer una idea, pero que no sepa que la posee). Sin embargo, este notar no nos provee la garantía de una certeza epistémica, a menos que las ideas notadas cumplan con un criterio de certeza otorgado por los requisitos de *claridad* y *distinción* (Descartes, 1951, I, XLV, pág. 21). Si se cumple con este criterio de certeza, se puede producir un juicio cierto en el que se muestre la ligazón entre el contenido de una vivencia y un hecho externo.

Si una idea sólo existe (y completamente) mientras es vivenciada, se requiere de un criterio de identidad para reconocer una idea como siendo la misma (o diferente de la) que se percibió anteriormente. ¿Dónde reside la relación de identidad de una idea, en el sentido en que un sujeto pueda decir “E es la misma idea que tuve ayer”, o “la idea es una idea- ejemplar del mismo tipo que E”? De algún modo, parece necesario marcar las ideas para reconocerlas como idénticas, o como ejemplificaciones de una misma *idea – tipo*. Para Descartes, la luz natural (facultad otorgada por Dios) hace todo el trabajo (1951, I, § XXX, pág. 17). Es la propia claridad y distinción de una idea la cual, iluminada por la luz natural, dibuja para la mente los contornos, las similitudes y las

diferencias de una idea con otras ideas. No hay posibilidad de error mientras el espíritu se atenga a lo inmediatamente notado, sin ulteriores juicios acerca de ese notar. El error puede provenir de diversas fuentes: los prejuicios (1951, I, § LXXI, LXXII, págs. 32 - 4), la fatiga (I, § LXXIII, pág. 34) y la ligazón de los conceptos a palabras que no corresponden a las cosas (I, § LXXIV, págs. 34 - 5). Por esta última razón, el lenguaje se convierte en fuente de errores y resulta conveniente apartarse lo más posible de las palabras para atender a las ideas. A pesar de ello, el lenguaje es necesario para realizar juicios y comunicar ideas.

A diferencia de Descartes, para Thomas Hobbes el lenguaje es esencial para realizar la tarea de identificación y reidentificación de ideas. Por eso, el lenguaje funciona como una marca, y los nombres son “postes indicadores instalados en un país desconocido” (1979, pág. 134): a pesar de su evidencia y su inmediatez, el flujo continuo e indetenible de las ideas puede convertirse en una sucesión caótica de vivencias sin sentido, a menos que podamos colocar ciertos “postes”: nombres comunes que relacionan y ordenan a las ideas. El concepto nace de la posibilidad de encontrar una relación entre una vivencia y otra o de entender a una vivencia como un *ejemplar* de un determinado *tipo*. Esta capacidad, propia de la luz natural de la razón, sólo puede realizarse plenamente cuando se posee un lenguaje. Para Thomas Hobbes, la posesión del lenguaje permite recordar las relaciones entre las ideas, relaciones que hemos descubierto mediante un trabajo inferencial. “El primer uso de los nombres es el de servir como *marcas* o *notas* de rememoración” (1979, cap. IV, págs. 139 -40), puesto que “nuestros pensamientos (son) propensos a deslizarse fuera de la memoria y forzarnos a un nuevo trabajo” (1979, pág. 140). La ausencia del nombre hace que las inferencias que hemos extraído se pierdan, pues la memoria sólo retiene un proceso inferencial si puede darle nombre. En ausencia de los nombres, las inferencias se presentarían como una ciega sucesión de ideas, sin que se pueda saber cómo se ha llegado a una última idea a partir de otra. La situación sería comparable a la de un barquero que navegase por diversas corrientes y que de pronto se encontrase en el mar, sin saber por qué ríos ha viajado ni cómo se conecta ese mar con los ríos anteriores. El lenguaje ilumina las conexiones entre las ideas. Razonar y utilizar un lenguaje son un mismo proceso, pues no puede hacerse lo primero sin lo segundo.

Gracias al lenguaje, “la consecuencia encontrada en un caso particular se registra y recuerda como una regla universal (...) Un idiota de nacimiento incapaz de retener en la memoria el orden de términos numéricos como *uno, dos y tres*, puede observar cada campanada del reloj y asentir a ella, o decir *una, una, una*, pero jamás sabrá qué hora está marcando” (Hobbes, 1979, pág. 142). Las concatenaciones de nombres y las de números son posibles gracias a la utilización de una *misma* facultad. “Podemos definir lo que resulta significado mediante la palabra *razón* cuando la incluimos entre las facultades de la mente. Pues la razón, en este sentido, no es sino cálculo (esto es, adición y sustracción) de las consecuencias de nombres generales convenidos para *caracterizar* y *significar* nuestros pensamientos; digo *caracterizar* cuando calculamos por nosotros mismos, y *significar* cuando demostramos o probamos nuestros cálculos a otros hombres.” (Hobbes, 1979, pág. 149)

La capacidad de realizar operaciones, propia de la razón, se instancia mediante las inferencias matemáticas y mediante las concatenaciones de lenguaje: ambos procesos son contrararias de una misma facultad, la razón, cuya tarea consiste en encontrar las relaciones entre diversas ideas. La razón tiene una doble tarea: dar forma a los pensamientos (*caracterizar*) y asignarles términos (*significar*) para que puedan ser comunicados. La tarea de *significar* es, sin embargo, subsidiaria de la *caracterización*. El lenguaje y las operaciones matemáticas heredan esta doble tarea de la razón. Es necesario que haya nombre para *caracterizar* ciertas ideas que sin el nombre no podrían ser pensadas con claridad y distinción. Este nuevo nombre genera un concepto, y por lo tanto, una nueva idea. Por eso podemos formar un nuevo nombre para recordar la idea formada por el nombre original. Así podemos proceder hasta el infinito. Pero tras cada grado en la nominación se pierde contacto con la concatenación de ideas que dio origen al primer nombre. Y no es imposible que se le otorgue un nuevo nombre a una serie de relaciones que no sólo jamás se dieron en las ideas, sino que tampoco podrían darse de ninguna manera: “El primer autor del lenguaje fue el propio *Dios*, que instruyó a Adán en la denominación de las criaturas por él presentadas a su vista (...). Pero fue suficiente para llevarle a añadir más nombres a medida que iban dándole ocasión la experiencia y el uso de las criaturas, y para unir las gradualmente a fin de hacerse comprender; y así, con el paso del tiempo,

fue consiguiendo el hombre tanto lenguaje como cosas a designar, aunque no tan copioso como el requerido para un orador o un filósofo” (1979, pág. 139). Los filósofos han agregado términos como “entidad”, “intencionalidad” y “quiddidad”, los cuales son “palabras sin sentido de la escolástica” (1979, pág. 139). Hobbes establece la imperiosa necesidad de las definiciones: el análisis preciso de un término de manera tal que se revele la concatenación de ideas que el nombre pretende subsumir. De lo contrario, el hombre “se verá enzarzado en un maraña de palabras como el pájaro en un cepo, y cuanto más luce más atrapado se verá” (1979, pág. 143). Ahora bien, ¿cuál es el peligro de esta maraña de palabras, y en qué consiste exactamente? Uno de los abusos del lenguaje que destaca Hobbes es el *mal registro de los pensamientos*. Poner inadecuadamente una marca es causa de que el nombre no signifique lo que se espera que signifique. Presumiblemente, este *mal registro* puede tener dos raíces: o bien, porque las ideas mismas son muy diferentes entre sí, y se pretende que el nombre común, haciendo caso omiso de esa diferencia, pueda subsumir esa heterogénea variedad de ideas en una misma característica universal, o bien porque el nombre parece aludir a otra idea diferente de la que pretende nombrar. Pero existe un tercer peligro en los nombres: cuando no nombran una idea y, por lo tanto, aquello que nombran no puede ser siquiera conceptualizable. Ejemplos de esto son los términos como “entidad” o “quiddidad” de la escolástica (1979, pág. 139 y 146).

Hobbes enfatiza la fuerte dependencia de las ideas discursivas con el lenguaje, de modo que no es posible vivenciar ciertas ideas si no se posee un soporte lingüístico en el que tales ideas puedan ser moldeadas y recordadas. Este modo de entender la vivencia asociada al lenguaje no estaba presente en Descartes, para quien en principio es posible separar de manera tajante el lenguaje con el que se realiza el juicio, y la idea⁴. Esto no significa que para Hobbes la idea dependa fundamentalmente de la palabra; en un sentido existe un “discurso mental” no lingüístico, que se realiza sin el auxilio del lenguaje público y a través del cual ocurren las inferencias discursivas. El lenguaje sólo es un soporte para moldear y recordar el discurso mental, pero las ideas

⁴ “[E] error principal y más común que pueda encontrarse consiste en que juzgo que las ideas que están en mí son semejantes o conformes a las cosas que están fuera de mí; pues ciertamente, si sólo considerara las ideas como ciertos modos o maneras de mi pensamiento, sin quererlas referir a algo exterior, apenas podrían darme ocasión de equivocarme” (Descartes, 1977, pág. 236)

no necesitan “discurrir en el lenguaje”; ellas tienen su propio modo de discurrir en la mente.

No hay dudas, sin embargo, de que el objeto primario del lenguaje, tanto en Descartes como en Hobbes, es la idea y no el mundo exterior. Pero la palabra *idea* tal como la utilizan estos pensadores y otros (Locke, Berkeley, Hume) se ve afectada de una enorme polisemia, y por ello algunos comentaristas contemporáneos declaran su desesperanza por encontrar una definición útil al respecto. Geoffrey Warnock (1953) declara al respecto de la noción de idea en Locke:

En este tema central, resulta imposible obtener de Locke una respuesta clara. Con el término “idea” significó, según su propia confesión, casi todo lo que uno guste elegir; nos dice que este término se refiere a “todo lo que sea objeto del entendimiento cuando un hombre piensa”, que significa lo mismo que ‘imagen (*phantasm*)’, noción, dato sensible (*species*), y también que hace referencia, en general, a objetos de la percepción. (Warnock, 1953. En Hacking, 1979. pág. 41)

En términos parecidos se pronuncia David Armstrong (1965):

Locke usa la palabra “idea” de una manera extraordinariamente amplia: al menos abarca:

- Percepciones sensoriales.
- Sensaciones corporales (cosas tales como dolores o cosquilleos)
- Imágenes mentales;
- Pensamientos y conceptos.

El uso de una sola palabra para abarcar esta clase heterogénea de cosas conduce a Locke a todo tipo de errores. (Armstrong, 1965 en Hacking, 1979, pág. 41)

Jonathan Bennett (1988) considera que esta polisemia puede deberse a que Locke utiliza la palabra “idea” para referirse, simultáneamente, a hechos mentales y a impresiones sensibles:

Un rasgo esencial en la teoría de Locke... (es) el uso de “idea” para referir a algo mental que podría dotar de significado a una palabra y referir a una impresión sensible o a un dato sensible. (Bennett, 1988, pág. 16)

Es esta doble referencia la que probablemente genere la polisemia advertida por Warnock, Armstrong y Hacking⁵. La formulación de esta doble referencia por parte de Bennett parece implicar que una impresión sensible no es “algo mental”, o no en el mismo sentido en que lo es aquello que otorga el significado a una palabra. En todo caso, el tratamiento por parte de Locke de ambos tipos de referencia a través de una sola palabra puede justificarse en tanto ambos pretenden cubrir exhaustivamente los requisitos de las vivencias enumerados más arriba (Privacidad, transparencia, irreducibilidad e incorregibilidad). ‘Idea’, entonces, es sinónimo de ‘vivencia’. Todo fenómeno de conciencia es esencial y necesariamente vivencial, y todo lo que no es vivencial no es fenómeno de conciencia o, lo que es lo mismo, no es una idea en la mente de alguien. ¿A qué se refiere, entonces, Bennett cuando habla de “datos sensibles”, por oposición a “algo mental” que dota de significado a una palabra? Si los datos sensibles pueden separarse del contenido mental, entonces difícilmente podemos decir que los datos son ideas. La mejor manera de interpretar esta cita es que en ambos casos se trata de “algo mental”, con la diferencia de que en lo que hace a los hechos mentales que Bennett llama “datos sensibles”, es posible encontrar un componente causal intencionalmente dirigido hacia un ámbito diferente al del propio ámbito mental. Pero, ¿qué es, específicamente, un “dato sensible” en este sentido? El componente causal del dato “señala” (gracias a su carácter intencional) su causa, la cual no puede ser un contenido de conciencia. No hay garantías de que esa dirección intencional apunte a una causa verdadera; tampoco hay garantía de (y de hecho no puede esperarse) que la causa sea semejante⁶ al dato sensible. En esta cadena causal (de la cual el dato sensible es el último eslabón, el único plena y cabalmente consciente) podemos encontrar los siguientes elementos:

- La realidad exterior (Propiedades y esencias reales).
- El estímulo.

⁵ En la cita de Warnock transcrita más arriba, podemos ver que a pesar de la aparente heterogeneidad de los conceptos involucrados en la definición de “idea”, es posible clasificar a todos ellos en estos dos usos que implican dos tipos de referencia diferente. ‘Imagen (phantasm)’ y ‘noción’ están usados como “algo mental que puede dotar de significado a una palabra” (así como los puntos c y d de la cita de Armstrong), mientras que ‘dato sensible (*species*)’ y ‘objetos de la percepción’ (así como los puntos a y b de la cita de Armstrong) pueden clasificarse de acuerdo a la definición de ‘idea’ como signo de una impresión sensible.

⁶La noción de “semejanza”, en lo que hace a la relación entre las ideas y el mundo exterior se tratará más adelante, en § 1.5.3

- El percepto.
- El dato sensible.

Según Jesús Mosterín (1981), los perceptos son:

patrones o plantillas de nuestro sistema neurosensorial, que nos permiten identificar formas perceptuales cada vez que se presentan en el continuo de nuestras sensaciones. (Mosterín, 1981, págs. 12-13)

Un percepto es el resultado inmediato de una percepción y sus consecuencias no lingüísticas posteriores. Una de las características del percepto es su relativa estabilidad: si bien los estímulos –que provocan el percepto– son informes e irregulares, el resultado de esos estímulos no lo es. A partir del percepto se conforma el dato. El dato es ya una construcción lingüística; es producto de una hermenéutica aplicada al percepto y, por lo tanto, de las funciones superiores del cerebro.

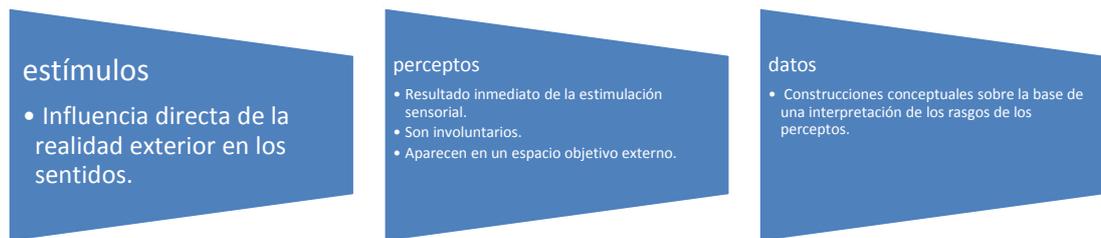


Figura 1 – Diferencias entre los estímulos, los perceptos y los datos.

El percepto es la consecuencia inmediata de la activación de una determinada modalidad sensorial. Por ejemplo, en lo que hace al sentido de la vista, los fotorreceptores (los bastoncitos y conos) son impresionados por determinados rangos de luz. Un estímulo luminoso cuyo rango esté entre los 495 y los 570nm produce un percepto que es identificado fenoménicamente como el dato sensible de “verde”. Es importante destacar que el percepto y la idea no son sinónimos: el percepto no es todavía consciente, pero es la *causa* de que se tengan determinadas experiencias sensoriales conscientes. En este sentido, las células fotorreceptoras ya son en sí mismas un sistema de clasificación y de selección de estímulos. A partir de la estructura funcional de la modalidad sensorial involucrada, es posible predecir determinadas configuraciones perceptuales. Si mismos rangos de luz impresionan al

ojo siempre de una manera distinta y azarosa, entonces el percepto conformado se encontraría en la misma situación que el estímulo: no permitiría la identificación de patrones perceptuales distintivos y el dato sensible fenoménico sería sumamente impredecible. Es esperable que, ante la invariancia de propiedades externas y estímulos, el percepto se comportara de manera correlativamente invariante y que, por ello, el dato sensible fenoménico también lo sea.⁷

¿Consideraría Locke que un percepto es también una idea? Es posible que el componente causal involucrado en el dato sensible incluya al percepto: la “huella causal” es la forma regular y reconocible de la modalidad del dato perceptivo: “ver algo rojo” es, esencialmente, reconocer en el rojo visto la “forma-de-la-rojez” como resultado de un estímulo externo. De este modo, que podamos reconocer un componente causal en un dato perceptivo (y no meramente que “tengamos” el dato como “la idea de algo rojo”, sin referencia a un hecho externo) implica que las huellas causales del percepto han de ser de algún modo esenciales y constitutivas del dato perceptivo.

Dijimos que Locke utilizaba la palabra ‘idea’ en una multitud de sentidos, que podían reducirse a dos tipos de ellos: las “ideas de datos sensibles” y las “ideas en la mente”. Estos dos tipos de ideas pueden distinguirse de acuerdo a su *dirección causal*: en las ideas del primer tipo existe un percepto que indica la huella causal de un estímulo, mientras que en las del segundo no. Sin embargo, para Locke toda idea tiene su origen en la percepción, de modo que incluso aquellas ideas que no nos “hablan” de un hecho externo, deben poseer de algún modo una conexión causal con el mundo. En Locke hay una noción que permite hablar acerca de esta conexión causal: *cualidad*. Las cualidades primarias y secundarias se infieren a partir del contenido de la idea, aunque ellas mismas no son un contenido de conciencia. Sin embargo, Locke es ambiguo con respecto a la función de una clase de cualidades, las cualidades secundarias: a veces

⁷ Incluso el dato fenoménico posee una estabilidad que los pasos anteriores podrían no poseer: tal es el caso de la constancia del color, fenómeno por el cual la percepción subjetiva del color se mantiene relativamente constante a pesar de los cambios en la iluminación. Cf.: Evan Thompson, Adrian Palacios, Francisco J. Varela, 1998.

habla como si tales cualidades fueran, precisamente, ideas⁸. Otras veces, en cambio, dice que son *potencias* de producir ciertas ideas⁹. A su vez, las cualidades secundarias son tipos de cualidades primarias imperfectamente conocidas. De modo que la cualidad secundaria amenaza con ser un híbrido entre componentes puramente fenoménicos y componentes causales externos. Estas ambigüedades tendrán sus consecuencias tanto en lo que respecta a la noción de ‘idea’ como en la noción de ‘cualidad’.

Señalamos al principio la “disunidad” de la idea. En vistas de lo dicho en el párrafo anterior, podemos adelantar que la idea es ontológicamente “desunida” en tanto se la concibe ambiguamente: a) como una entidad *puramente* mental y b) como una entidad mental que *sin embargo tiene –como parte de su esencia- una dirección causal del mundo hacia la mente*. Desde el punto de vista de a), la idea solamente es el dato sensible fenoménico. Desde el punto de vista b) la idea incluye –como una huella causal- elementos del percepto, del estímulo y de las propiedades del mundo exterior. El término idea es *polisémico*, pero la ontología de la idea, en tanto que fenómeno, es *desunida* puesto que cabalga ambiguamente entre una existencia puramente mental y una existencia como propiedad del mundo exterior.

1.2. Cualidades.

En el marco cartesiano para el cual existen dos tipos de sustancias -una mental y otra extensa-, se corre el riesgo de que la explicación del mundo no pueda unificarse si ambas sustancias no pueden explicarse reductivamente. La *idea* es el contenido propio de la sustancia mental, y en algunas de las ideas parece haber una huella del mundo exterior, lo cual en principio podría servir para tender el puente entre los hechos internos y los externos. Estas huellas nos señalan una dirección causal, y parecen indicarnos la existencia de ciertas cualidades que dejan su impronta en la idea. Es importante destacar, sin embargo, que sólo pueden vivenciarse propias ideas y nunca

⁸ “[Las cualidades secundarias son] ideas tales que no son más que otras relaciones con otras sustancias y, consideradas tan solo por sí mismas, no están realmente en el oro”. (Locke, [1710] 1956, II, XXIII, § 37, pág. 299)

⁹ “La amarillez no está realmente en el oro, pero es una potencia en el oro para producir esa idea en nosotros cuando se sitúa en la luz debida” ([1710] 1956, II, XXIII, § 10, pág. 281)

se vivencia un acontecimiento externo: la idea, aunque contuviera una huella causal de la exterioridad, es ella misma totalmente interna.

El pensamiento de Descartes, Locke, Berkeley y Hume se caracteriza por el *internismo semántico*¹⁰: el lenguaje se refiere, en primera instancia, a las ideas, y sólo derivadamente al *mundo exterior*. La noción de internismo semántico está incluida en la de internismo intencional: todo acto mental (no sólo lingüístico) está, en primera instancia, dirigido hacia las propias ideas, y sólo de manera inferencial, derivada, dirigido al exterior. De esta situación epistémica de pensadores como Locke y Descartes, Hilary Putnam (1984, pág. 137) extrae una consecuencia desconcertante a la que denomina “solipsismo metodológico”: los pensadores que están dentro de este marco cartesiano no pueden explicar un fenómeno tan familiar como los celos: que una persona X esté celosa de otra persona Y se explica, en este marco internista, como el hecho de que X está celoso del contenido de sus propias ideas, las cuales (faliblemente) representan a Y. En otras palabras: uno puede tener la certeza de que está celoso de una idea, aunque jamás puede saber si dicha idea representa a una persona Y que constituye la intención terminal de esos celos. Dado que el marco cartesiano restringe severamente la garantía de acceso a la exterioridad, es necesario contar con herramientas conceptuales que nos indiquen cuándo estamos en presencia de una idea cuyo contenido nos instruye acerca de *lo que pasa afuera*, y cuándo no. Todo proceso intencional está dirigido primariamente hacia la interioridad, pero de algunas ideas es posible obtener información acerca de su origen, y en ese sentido su contenido nos da indicios acerca de algo más que esa misma idea. Los indicios son puramente inferenciales; no hay más que una serie de relaciones falibles (pero hasta cierto punto fiables) que permiten producir juicios acerca de *acaecimientos* y no sólo acerca de *vivencias*. Si hablamos acerca de nuestras ideas, basta con que nos refiramos a ellas en términos de *vivencia*, pero cuando pretendemos dar el salto judicativo acerca de su relación con un acaecimiento utilizamos la expresión *percepción*. Así, las ideas se vivencian, pero los acaecimientos se *perciben*. El uso del verbo “percibir” implica el éxito en la inferencia, a partir de una idea puramente vivenciada: no se

¹⁰ “Diremos entonces que una teoría *internista* de la representación es una según la cual el objeto intencional de todos los estados intencionales es *inmanente*, y una teoría *externista* es una según la cual el objeto de algunos es *trascendente*”. (García Carpintero, 1996, pág. 79)

puede decir que se percibe un objeto si el objeto no existe (o si posee cualidades diferentes a las vivenciadas), aun cuando se pudiera estar vivenciando ciertas cualidades objetivas. Pero la garantía de que una vivencia sea, además, una percepción, es problemática. La situación del internista semántico es comparable a la de una persona que está encerrada en una habitación sin ventanas: el único modo de saber qué pasa afuera consiste en inspeccionar el contenido de alguna parte de la habitación que le dé un indicio acerca de lo que pasa afuera. Si en la habitación hubiera un monitor conectado a una cámara, la cual estuviera conectada (presumiblemente) hacia el paisaje exterior, la sola inspección de la cámara le daría información indirecta acerca de lo que pasa afuera. Desde luego, este es un indicador falible: el monitor podría no estar conectado a la cámara, o podría estar transmitiendo imágenes de otro lugar. Un *operador de cable maligno* podría representarle imágenes de paisajes inexistentes, o incluso podría representar algo existente de un modo sistemáticamente incompatible con lo que efectivamente existe. ¿Cómo hacer para garantizar el vínculo entre el monitor y el mundo exterior? ¿Cómo vincular necesariamente el contenido de una idea con la impresión externa que la ha originado? O, para extremar el ejemplo considerado, ¿cómo salir de la habitación *sin salir de la habitación*?

Descartes propone examinar el contenido de las ideas para comparar los grados de realidad objetiva que representan. La realidad objetiva es el contenido de una idea; la realidad formal es el hecho (real) que fue causa de dicho contenido (Descartes, 1951, § XVII y XVIII; 1977, meditación III). La argumentación es la siguiente: si el sujeto poseyera una idea cuya realidad objetiva fuese superior a la propia realidad formal del sujeto, entonces se habrá encontrado una idea cuya realidad formal no puede ser el propio sujeto. La idea de Dios incluye en su realidad objetiva la noción de un ser que, por definición, es infinitamente superior a la propia realidad formal del sujeto (suponiendo que hablamos de un sujeto humano y no del mismo Dios). Por lo tanto, la realidad objetiva de dicha idea no pudo haber sido creada por un alma humana. Dada la desproporción entre el contenido de la idea de Dios y la pobre eficacia causal del sujeto humano, es necesario postular un ente con tanta eficacia causal como contenido tenga la idea (Recordemos que no es posible dudar acerca de la propia idea,

ni de su contenido: el notar es esencialmente incorregible). De este modo, se prueba que existe Dios y por lo tanto el contenido de ciertas ideas (al menos de una de ellas) puede informarnos necesariamente acerca de un hecho externo a las propias ideas y a la propia subjetividad. De este *excursus* inferencial, Descartes deriva la negación de un hipotético genio maligno que pudiera interferir con la verosimilitud del contenido de las ideas, en tanto semejantes a hechos del mundo externo. Probada la existencia de Dios, el resto del mundo sólo depende de una propiedad divina: la veracidad. Dios es veraz; esto es: su propia existencia garantiza la posibilidad del conocimiento (1951, I, §§ XXIX, XXX, XXXI, IV, CCVI). El error es posible, pero se explica por las deficientes facultades humanas (prejuicios, fatiga, conceptos erróneos) y no por un desacople intrínseco (y malintencionado) entre las ideas y el mundo externo (1951, I, §§ XXIX, XXXI, XXXVI, XXXVIII).

Una vez que se ha probado la conexión entre el contenido de una idea y la realidad formal que le da origen, no se ha hecho todo el trabajo. Todavía se requiere de un criterio para reconocer cuándo el contenido de una idea posee una conexión con el mundo externo, y cuándo no. La alucinación completa es posible, ya no por la intervención de un genio maligno, sino por el total mal funcionamiento de una estructura cognitiva ya no (sólo) debida a los prejuicios, a la fatiga o a los errores de lenguaje, sino incluso a alguna inadvertida enfermedad mental generalizada. ¿Qué ocurre en un estado alucinatorio? Tengo –por ejemplo- la idea de una mujer que camina por la calle desierta (ese es el contenido de mi idea) pero, aunque Dios me garantiza la posibilidad de la conexión entre el contenido de mis ideas y hechos externos correspondientes a ellas, no queda por ello garantizado el éxito de dicha conexión. De hecho (aunque yo no lo sé; presuntamente lo sabría alguien que estuviera observando el mismo paisaje que yo, o el mismo Dios) no hay una mujer caminando por la calle desierta; no hay indicios epistémicos en mi experiencia que pudieran garantizarme la verdad del contenido de la idea: en lo que a mí respecta (incluso con el auxilio del criterio de certeza, la claridad y la distinción), el contenido de la idea parece haber sido causado por una realidad formal equivalente a la idea. Dado que toda idea con contenido perceptivo puede ser ficticia (porque yo mismo podría ser el creador de dicha idea), quizás la realidad formal de la idea “una mujer cruzando la

calle desierta” soy yo mismo. Necesitamos encontrar un correlato causal entre la idea y el hecho que provoca la ida. Dichos candidatos son las *cualidades*. Una cualidad es, en su acepción más general, todo aquello que no es una idea pero que sirve de causa para la vivencia de dicha idea. Gracias a las cualidades podemos tender los puentes intencionales que nos permiten hablar no sólo de nuestras vivencias, sino también de nuestras percepciones. El concepto de cualidad es estrictamente *relacional*: la cualidad nos permite tender puentes entre el mundo externo y la interioridad de la idea. Algunos aspectos de la realidad formal de un objeto¹¹ impactan en la estructura cognoscitiva del sujeto y se convierten en ideas. La postulación de estas cualidades, como intermediarios entre lo objetivo y lo subjetivo, es característica del *realismo por representación* (García Carpintero, 1996, pág. 77). El realismo por representación consiste en las dos afirmaciones siguientes:

- 1) Los objetos intencionales inmediatos de los estados mentales no son objetos reales ni sus propiedades, sino entidades mentales (ideas)
- 2) Estos objetos mentales representan en virtud de *relaciones causales* a los objetos de la realidad y sus propiedades (García Carpintero, 1996, pág. 74).

Mediante una parábola, Descartes pretende ilustrar la naturaleza de nuestros estados mentales dirigidos a los acaecimientos del mundo exterior (García Carpintero, 1996, págs. 77-8):

Sin duda habéis visto la necesidad de utilizar un bastón para guiaros cuando caminabais sin luz por lugares difíciles por la noche. Así mismo, os habréis percatado de que mediante el extremo del bastón podéis apreciar la existencia de diversos objetos que se encuentran a vuestro alrededor, e incluso que podéis distinguir si son árboles, piedras, arena, agua, hierba, barro o algún otro objeto semejante. Verdad es que esta forma de sentir es un poco oscura y confusa para aquellos que no han tenido una gran práctica. Pero si consideráis el constante ejercicio de aquellos que, habiendo nacido ciegos, se han servido de tal medio durante toda su vida, entonces encontraréis tan perfecta y tan exacta que podríamos afirmar que ven con sus manos o que su bastón es el órgano de un sexto sentido, que les ha sido dado al carecer de la vista. Para establecer una comparación a partir de esto, deseo que penséis que la luz no es otra cosa en los cuerpos, que son llamados luminosos, que un cierto movimiento o una acción muy rápida y muy viva que se dirige a nuestros ojos a través del aire y de los otros cuerpos transparentes, de igual forma que el movimiento o la resistencia de los

¹¹ Uso “objeto” en un sentido amplio para referirme a cualquier hecho, acaecimiento o suceso externo.

cuerpos que encuentra este ciego llega a su mano a través del bastón. Tal consideración os impedirá encontrar extraño (...) que por medio de la luz podamos ver toda clase de colores, así como que estos colores no sean otra cosa en los cuerpos, sino las diversas formas en que los mismos reciben y reflejan la luz contra nuestros ojos, si consideráis que las diferencias constatadas por un ciego entre diversos árboles, piedras, agua y cosas semejantes por medio de su bastón no le parecen menores de lo que son para nosotros aquellas que existen entre el rojo, el amarillo, el verde y todos los otros colores. Y sin embargo todas aquellas diferencias no son otra cosa en todos esos cuerpos que las diversas formas de mover ese bastón o de resistir a sus movimientos. (...) **No es necesario suponer que haya nada en estos objetos que sea semejante a las ideas o sentimientos que de ellos tenemos**, de la misma forma que (...) la resistencia o movimientos de esos cuerpos, que es la única causa de los sentimientos que (el ciego) tiene, no es en nada semejante a las ideas que concibe (Descartes, René, *Dióptrica*. En *Discurso del Método, Dióptrica, Meteoros y Geometría*, Madrid, Alfaguara, 1981, pp 61-62. Negritas añadidas)

Hemos resaltado en negrita un aspecto crucial de cierto tipo de cualidades: la no semejanza entre la cualidad que causa la idea, y la propia idea. He aquí entonces un hecho desconcertante: necesitamos encontrar un correlato externo que posea el mismo grado de realidad formal que la realidad objetiva contenida en la idea, y necesitamos corroborar de algún modo que esa realidad formal es diferente de mí mismo. Eso parece embarcarnos en la búsqueda de una cualidad externa causalmente suficiente, y tal vez modalmente similar al *producto* de dicha cualidad, esto es, la idea. En otras palabras: si vivencio la idea de una mujer cruzando la calle desierta, tal idea sólo puede estar causalmente justificada si efectivamente afuera hay algo que posea un contenido causal suficiente para generar la idea que estoy teniendo. Si hay una disparidad demasiado grande, entonces o bien estoy alucinando, o bien mi mecanismo perceptivo no está preparado para comunicarme “lo que pasa afuera”¹². Pero ahora Descartes nos adelanta que no es necesario suponer que haya algo semejante a nuestras ideas en el mundo externo. Por lo tanto, o bien nuestras ideas no son semejantes a su causa (y en ese caso, estamos en una situación epistémica no muy diferente a la que teníamos cuando postulábamos la existencia del genio maligno), o bien debemos abandonar la noción de que nuestras ideas representen fielmente un hecho externo. En otras palabras: en este *realismo por representación*, o

¹²Este es el argumento conocido como “el velo de la percepción”.

abandonamos el *realismo* (pues las ideas no dan indicaciones fidedignas acerca de la realidad exterior), o abandonamos la *representación* (en el sentido de que una idea “representa” un hecho externo).

La estrategia del examen del contenido representacional, por contraste con la realidad formal de la causa, resulta insatisfactoria. Con esta estrategia (y siempre que estemos de acuerdo en aceptar el uso de la noción aristotélica de causa eficiente para la totalidad de los contenidos de conciencia) sólo podremos probar con certeza la existencia de Dios, teniendo previamente como cierta e indubitable la propia existencia. Pero, exceptuando la existencia de Dios, no tenemos medios seguros para conocer fidedignamente los hechos externos *en tanto que* hechos externos: la única conexión epistémica entre el yo y el mundo es la veracidad divina. Ahora bien, para que la veracidad tenga un valor epistémico (y no se convierta en la mera probabilidad no garantizada de que un dato perceptivo se corresponda con un hecho objetivo, dado que pueden mediar limitaciones propias de mi sistema cognitivo que no son producto de un engaño divino) se requiere de un método para distinguir entre hechos puros de conciencia y hechos de conciencia que contienen un *percepto*, esto es, información fenoménica de los patrones generales de estimulaciones que contienen una huella causal. Las palancas epistémicas son las nociones de *cualidades primarias* y *cualidades secundarias*. En los *Principios de Filosofía*, Descartes enumera la diferencia en los modos de conocer en lo que hace a los cuerpos, magnitudes, figuras y movimientos por un lado, y los colores, sabores, sonidos y aromas (1951, I, § LXIX): los primeros son percibidos clara y distintamente, mientras que los segundos sólo mantienen ese estatus si *no* los dirigimos intencionalmente hacia una supuesta exterioridad (1951, I, § LXVI): nos equivocamos al dirigirnos intencionalmente hacia alguna parte de nuestro cuerpo cuando formulamos el juicio de que algo nos duele. El dolor está enteramente en la sustancia pensante; cualquier acto ostensivo que indique una parte del cuerpo dolorida implica un error de categorías (pues la sustancia extensa no tiene al dolor entre sus propiedades). Sin embargo, el dolor al igual que la ira, el amor y el odio, son “pensamientos confusos que la mente no tiene por sí sola, sino del hecho de que sufre alguna parte del cuerpo al que está íntimamente unida” (Descartes, 1951, IV, § CXC). Es curioso que Descartes diga que es *el cuerpo* el que *sufre* (y no el alma dolorida).

Queda claro, sin embargo, que atribuye (del mismo modo que lo hará Locke) la adquisición de los sabores, olores, sonidos, colores y texturas a la imperfección de los sentidos para capturar las cualidades propias del objeto percibido, pues todo lo que hay en la sustancia extensa son “formas, situaciones, tamaños y movimientos de sus partes”(1951, IV, § CXCVIII, pág. 236) y no colores, gustos, sonidos, calor o frialdad. De hecho, no poseemos los sentidos para saber de qué están hechas las cosas, sino para sobrevivir en el mundo:

(...) Las percepciones de los sentidos no se refieren sino a esta conjunción del cuerpo humano con la mente, y que nos muestran ordinariamente en qué pueden serle útiles o dañinos los cuerpos externos, pero no nos enseñan, sino a veces y por accidente, cómo existan en sí mismos. (1951, II, § III, pág. 38)

En rigor, los tipos de cualidades en Descartes necesitan de un apoyo adicional: la noción de “cualidad” se concibe como mediadora entre el aspecto interno y el externo, pero a su vez existen otros mediadores: los *nervios* y los *espíritus animales*. Los primeros son emisarios del hecho perceptivo y reciben la estimulación inmediata de la exterioridad *de la que son parte* (1951, IV, §§ CLXXXIX, CXC, CXCVI, CXCVIII). Los segundos, por su naturaleza espiritual, decodifican el estímulo para transformarlo en sensación. Dado que el alma está alojada en el cerebro (1951, IV, § CLXXXIX), es necesario explicar no sólo cómo el cerebro se relaciona con el resto del cuerpo, sino también cómo el alma se relaciona con el cerebro. De allí surge una división tripartita entre alma, cerebro y sentidos (o resto del cuerpo en general). “El alma no siente sino en tanto está en el cerebro” (1951, IV, § CXCVI, pág. 234); el cerebro se conecta con los sentidos a través de los nervios; los nervios informan del dolor o del placer y el alma, a través de los espíritus animales, como receptora terminal de un proceso nacido en la sensación, se informa acerca de esos hechos que nacen en los nervios y confluyen en el cerebro.

En Locke ([1710] 1956) encontramos explícitamente la caracterización de dos tipos de cualidades. Las cualidades primarias son para Locke,

Tales que son por completo inseparables del cuerpo, cualquiera sea el estado que éste se encuentre; y tales que las conserva en todas las alteraciones y cambios que dicho cuerpo pueda sufrir a causa de la mayor fuerza que pueda ejercerse sobre él. Esas cualidades son tales que los sentidos constantemente las encuentran en cada partícula

de materia con bulto suficiente para ser percibida, y tales que la mente las considera como inseparables de cada partícula de materia, aun cuando sean demasiado pequeñas para que nuestros sentidos puedan percibir las individualmente (Locke, [1710] 1956, II, VIII, 9, pág. 113)

Bennett comenta al respecto de esta cita que las cualidades primarias son *aquellas que un cuerpo no puede perder* (1988, pág. 122): si dividimos un cuerpo en varias partes, cada una de esas partes conservará las mismas cualidades que el cuerpo original. En este punto, entonces, las cualidades primarias son todas aquellas propiedades que adjudicamos como necesarias y suficientes para juzgar que algo es un *cuerpo*: solidez, extensión, figura, movimiento o reposo y número (Locke, [1710] 1956, II, VIII, § 9, pág. 112). Las cualidades secundarias, por el contrario, tienen una casi completa dependencia de la mente que percibe al cuerpo, y son propiedades *relacionales*, es decir, no son inherentes al cuerpo (al menos, de modo tan claro y directo) como presumiblemente lo son las cualidades primarias. Las cualidades secundarias son: color, temperatura, dureza /blandura, olor, sabor y sonido ([1710] 1956, II, VIII, § 10, págs. 112 - 3.)

1.2.1. Propiedades generales de la experiencia: tres requisitos que deben satisfacer las cualidades primarias.

En el primer apartado (§ 1.1) se analizó la noción de idea. La idea es desunida en tanto que, por una parte, es un hecho absolutamente mental, pero por otra parte contiene una huella causal del acaecimiento externo que le da origen. El hecho de que exista una huella causal es lo característico del *realismo por representación*: la idea representa un hecho externo en virtud de este indicio causal, y en ese sentido la propia idea contiene algo de la externidad que representa. En el apartado anterior (§ 1.2) establecimos que tanto la referencia de los términos conceptuales como la intencionalidad de los estados mentales están dirigidos, en primer término a las propias ideas y sólo de manera derivada a los hechos del mundo externo¹³. A esto lo

¹³ Estrictamente, el concepto fenomenológico clásico no denominaría “intencional” a una idea cuyo objeto es su propio contenido. La intencionalidad se caracteriza, precisamente, porque los objetos de la conciencia no son sus propios contenidos, sino las cosas en cuanto tales. Estamos tomando la noción de

habíamos llamado *internismo semántico e intencional*. Si todos los acaecimientos externos solo se derivan o se infieren de las propias ideas, es necesario contar con criterios para decidir cuándo efectivamente un hecho se deriva de algo exterior y cuándo se trata de una ficción o alucinación. La sola división en cualidades no es suficiente: se requiere de nociones adicionales para aceptar que una determinada idea es producto de un determinado tipo de cualidad -de otro modo, la división en dos tipos de cualidades sería arbitraria. En este apartado estableceremos algunos requisitos que podrían servirnos de criterios para distinguir los tipos de cualidades y, por lo tanto, para decidir si estamos ante un hecho perceptivo auténtico o si se trata de una alucinación. Si encontramos la línea que nos permita separar la experiencia real de la experiencia alucinatoria, habremos hallado a su vez un modo fiable de conexión entre nuestras ideas (la experiencia fenoménica consciente) y el mundo. El próximo paso consistirá en el intento de reducción de esa experiencia fenoménica a los acaecimientos mecánicos del mundo.

La división entre cualidades primarias y secundarias requiere de la aceptación de la noción de *cuerpo*. Para Descartes, un *cuerpo* es esencialmente una *confluencia singular de propiedades espaciales*: “La extensión en largo, ancho y profundidad que constituye el espacio, es absolutamente la misma que constituye el cuerpo. Pero la diferencia está en que la consideramos en el cuerpo como singular y juzgamos que cambia cuantas veces cambia el cuerpo” (Descartes, 1951, II § X). La extensión es el atributo fundamental del cuerpo (allí donde encontramos extensión, encontramos cuerpo), de modo que resulta contradictorio suponer que puedan existir espacios vacíos (1951: II, caps. XII, XVI). La expresión “espacio vacío” sólo puede utilizarse de modo relativo, en el sentido de que no contiene lo que esperaríamos, o sólo contiene aire (1951: II, caps. XVII). La extensión en el espacio es la única idea clara y distinta que podemos obtener acerca del mundo exterior. Esa claridad y distinción parece derivarse del hecho de que las cualidades primarias otorgan una estructura global y sistemática a la experiencia perceptiva. Gracias a las cualidades primarias, la experiencia posee *coherencia, continuidad y consistencia inferencial*. Como veremos, las cualidades secundarias no otorgan esas mismas características a la experiencia, con lo cual esta inherencia masiva

“intencional” en un sentido más general, en la cual los propios objetos de conciencia pueden dirigir intencionalidad hacia sus propios contenidos.

puede servirnos como guía para distinguir entre tipos de cualidades. Caractericemos estas tres propiedades generales de la experiencia otorgadas por las cualidades primarias:

- ***Coherencia***: la concurrencia de las cualidades de solidez, extensión, figura, movimiento o reposo y número otorgan a la experiencia una base para entenderlas como manifestaciones de un único hecho perceptivo. En otras palabras, percibimos a la multitud de esas cualidades como concurrentes en un mismo hecho (objeto, sustancia) y no como cualidades desconectadas o “huérfanas”.
- ***Continuidad***: Las cualidades primarias son estables; esto es: la modificación en el carácter perceptivo de una de ellas, conlleva a una modificación de la experiencia de la totalidad de estas cualidades. Por ejemplo, si a un cubo sólido se le redondean sus aristas hasta convertirlo en una esfera (es decir, si se cambia la cualidad de “figura”), se habrán modificado simultáneamente la solidez, extensión, movimiento o reposo y número en la percepción de esa figura.
- ***Consistencia inferencial o derivabilidad perceptiva***: dado un conjunto de cualidades primarias, podemos verídicamente realizar inferencias acerca de la totalidad de ese conjunto, y no meramente de cualidades desconectadas o huérfanas. Por ejemplo, determinando el grado de solidez, la extensión, la cantidad y la figura de un dato perceptivo, es posible inferir la extensión, la cantidad y la figura de otro dato perceptivo (que involucre la misma referencia desde otro punto de vista, más lejos, más cerca o mientras se mueve el sujeto percipiente) desde otro ángulo o bajo condiciones diferentes. García Carpintero pone esta propiedad en términos de “esperabilidad”:

Las relaciones entre las ideas espaciales nos llevan a esperar, pongamos por caso, que un cubo que contemplamos desde una cierta perspectiva presente tal y cual apariencia cuando se le hace rotar 30 grados. Haciéndolo rotar así, confirmamos la predicción y con ello la conjetura de que las relaciones espaciales objetivas tienen una estructura análoga a la que tienen las ideas espaciales que causan (sobre la cual hemos hecho la conjetura) (1996, pág. 143)

(De aquí en adelante nos referiremos a estas propiedades como “requisitos C-C-C” de la percepción de cualidades primarias)

Presumiblemente, las cualidades secundarias no tendrán estas tres características, y de hecho así parece que ocurre: en nuestro campo perceptivo, los sonidos, los colores, los sabores y las experiencias táctiles pertenecen a espacios modales diferentes que bien podrían no ser coherentes entre sí, ni en sí mismos (y de hecho muchas veces no lo son, como ocurre cuando ponemos sucesivamente nuestras manos en un balde de agua: con la mano izquierda percibimos tibio, con la mano derecha percibimos frío) (Locke, [1710] 1956, II, VIII, § 21, pág. 118); no son continuas (la modificación en la percepción de un sonido no implica la modificación en la percepción de un color); y no se puede hacer derivaciones fiables a partir de ellas (a veces un mínimo cambio de condiciones provoca un enorme cambio en la naturaleza fenoménica de la percepción, como cuando se disminuye levemente la luz de una habitación y los colores se vuelven notoriamente menos contrastantes)

Bennett propone dos casos imaginarios para enfatizar las diferencias entre cualidades primarias y secundarias: los casos de las *aberraciones perceptivas* (1988, pág. 130). Supongamos dos sujetos: El sujeto (C) ve dos cosas que tienen el mismo color cuando, de hecho, no lo tienen; y el sujeto (T) ve y siente dos cosas como siendo del mismo tamaño, cuando en realidad no lo tienen. (C) posee una aberración en la percepción de un tipo de cualidades secundarias. (T), en cambio, padece un mal que afecta su percepción de una clase de cualidades primarias.

“C se enfrenta a una cosa roja y a una cosa blanca y nos convence de que ambas tienen exactamente el mismo color. Cree nuestra afirmación de que tienen colores distintos y, puesto que no difieren de ninguna otra forma, si fuese necesario podríamos probarle que podemos ver en ellas una diferencia y él no. También C podría descubrir que los dos objetos reflejan luz de diferentes longitudes de onda y podría saber que las longitudes de onda corresponden, usualmente, con los colores vistos. Pero si ignora lo que dicen los demás acerca de los dos objetos e ignora hechos esotéricos de óptica, él podría no llegar a saber que tiene un defecto sensorial. *Una falla en la discriminación de las cualidades secundarias en alguien que, fuera de eso, es sensorialmente normal, puede pasar, y pasa en ocasiones, desapercibida a través de cualesquiera variaciones de distancia, ángulo de visión, condiciones de luz, etcétera* (Bennett, 1988, pág. 130, cursivas añadidas)

De acuerdo a la caracterización dada más arriba, una aberración en una cualidad secundaria determinada no debería arrastrar aberraciones en otros actos perceptivos. Esto significa que las cualidades secundarias no poseen el carácter de *continuidad*: la aberración en una cualidad secundaria no implica una modificación masiva en la totalidad de las cualidades secundarias concurrentes en una experiencia. A su vez, la experiencia no se vuelve *incoherente* por esta aberración (se pueden seguir considerando a las cualidades primarias asociadas como unívocamente referidas a un objeto, aunque este objeto se presente como verde en vez de rojo), ni deja de poseer *derivabilidad perceptiva*: aunque localmente el objeto se percibe de un color diferente, la totalidad de la experiencia permite realizar inferencias perceptivas fructíferas relativas al punto de vista del sujeto. Por ejemplo, si (C) camina alrededor del objeto que ve verde (y que, según testimonio de sus coespecíficos, debería ver rojo), su incapacidad para verlo del color “que en verdad posee”¹⁴ no limita su capacidad de predecir cómo irán apareciendo las cualidades primarias del objeto a medida que él se mueve. Para decirlo en pocas palabras, una aberración local en una idea derivada de cualidades secundarias no *contamina* a la *totalidad* de la experiencia, sino específicamente a la modalidad de esa percepción (en este caso, el color en la percepción visual)

Pero, ¿qué ocurre con (T)?

Contrastemos el anterior con el caso de T quien, apelando a lo que ve y siente, juzga que cierto jarro tiene el mismo tamaño que cierto vaso que es, en realidad, más bajo y más estrecho que el jarro... En este caso, podemos poner el vaso dentro del jarro; o llenar el jarro con agua y luego llenar el vaso con ésta y tirar el agua restante, o situar ambos recipientes en una mesa y llevar la mano de T por encima del borde del vaso hasta que choque con el jarro, etcétera. ¿Qué es lo que debemos suponer cuando T se enfrenta a estas manipulaciones de los dos objetos? Tan solo hay dos posibilidades pertinentes. (a) Cuando manipulamos el vaso y el jarro, T se da cuenta de lo que sucede y así, rápidamente se hace consciente de que estaba equivocado acerca de los tamaños relativos. (b) Cada vez que ideamos algún suceso con el vaso y el jarro, T lo percibe mal de tal manera que lo que ve y siente aun encaja tersamente con su juicio original acerca de los tamaños. (Bennett, 1988, pág. 131)

¹⁴Dejamos entrecomillada esta expresión en vista de las consideraciones que haremos más adelante (en la segunda parte) en torno a la fenomenología y ontología del color.

T posee una incapacidad local para distinguir tamaños. Pero si ocurriera la posibilidad (b), entonces su incapacidad se extendería masivamente hacia otros experimentos. Las pruebas cruciales que le proponíamos a T (introducir el vaso en el jarro, pasar la mano por los bordes de ambos cuando uno está dentro de otro, etcétera) nos llevarían a extender la aberración hacia cualquier situación perceptiva y no meramente a este caso particular. La aceptación de la posición b) nos lleva a la presuposición de que una incapacidad perceptiva en las cualidades primarias podría dar por tierra con las tres características que listamos más arriba: la percepción deja de ser masivamente coherente, continua e inferencialmente consistente.

Para obtener una analogía entre la “ceguera” de tamaño y la acromatopsia, debemos adoptar, entonces, la suposición b). Esto requiere que le concedamos a T incapacidades como las siguientes: que no puede ver o sentir que el vaso está dentro del jarro (o que el jarro no se ha ampliado, o que no se ha contraído el vaso); que no puede ver o sentir que el vaso está lleno de agua (o que queda agua en el jarro luego de que hemos llenado con éste el vaso); que no puede ver o sentir el borde del vaso (...) (Bennett, 1988, pág. 131)

Bennett arguye que una aberración en las cualidades primarias nos conduce a aceptar un conjunto de *aberraciones compensatorias*: si se ha fallado en notar que ha quedado agua en el jarro una vez que introducimos el vaso, también se fallará en notar ciertos hechos cuando el agua se utiliza en otras ocasiones (apagar una vela, arrojarla al rostro). “Lo mismo sucede con cualquiera de las otras aberraciones sensoriales con las que reforzamos la inicial: cada una requiere de nuevos refuerzos que, a su vez, exigen otros y así indefinidamente” (1988, pág. 132)

La posibilidad de estas aberraciones compensatorias para explicar una sola aberración persistente parece indicarnos que las cualidades primarias de una cosa están relacionadas entre sí, de modo que las cualidades primarias de un objeto están en estrecha relación con las cualidades primarias de otros objetos, en cualquier tiempo y cualquier espacio. Esta posibilidad nos indica que las experiencias relacionadas con las cualidades primarias cumplen con los requisitos C-C-C. La evidencia de su componente causal se refuerza precisamente por la persistencia en el cumplimiento de esos requisitos: resulta implausible esperar que una aberración infecte masivamente a toda experiencia de cualidades primarias de similar contenido a la experiencia aberrante,

pues cada percepción se convierte en una nueva prueba crucial C-C-C. Dado que las percepciones originadas en cualidades primarias se conectan entre sí de un modo congruente, es plausible suponer que el cumplimiento de los requisitos C-C-C no es una pura consecuencia de la estructura perceptiva del sujeto. Hay tantos tipos fiables y cotidianos de conexiones entre cualidades primarias que resulta por lo menos sospechoso suponer que tales conexiones son sistemáticamente gratuitas, inventadas o falsificadas. Entre estos casos fiables podemos mencionar todas las generalizaciones empíricas cotidianas: algunos pocos ejemplos son: “de dos cosas rígidas, la menor no puede contener a la mayor; una cosa no puede detener la caída de otra sin tocarla; un cubo no puede rodar sin saltos por una superficie plana; la marca de un disco circular sobre la cera será circular”. (Bennett, 1988, pág. 133)

Una vez que corroboramos sistemáticamente el cumplimiento de los requisitos C-C-C para las cualidades primarias, podemos estar en posición de hablar acerca de cuerpos y no meramente de ideas, pues este cumplimiento parece ser congruente no con el funcionamiento de nuestra mente, sino con algo independiente que de modo sistemático obliga a ese cumplimiento. La violación de los requisitos C-C-C nos deja un campo perceptivo no unificado y no congruente: de no cumplirse con esos requisitos, podríamos percibir determinadas ideas de cualidades primarias desde una perspectiva, pero no desde otra; o en dos instantes inmediatamente sucesivos y en una misma situación perceptiva podríamos percibir dos conjuntos de ideas de cualidades primarias diferentes y no unificados: sería imposible dar un sentido preciso a la noción de “cuerpo” cuando ciertos conjuntos de ideas perceptivas acerca de cualidades primarias no confluyesen en un todo congruente. Tomando la definición de cuerpo de Descartes (el cuerpo es una confluencia singular de propiedades espaciales), una aberración en la percepción de una idea obtenida a partir de cualidades primarias provocaría que ciertas propiedades espaciales no confluyeran, o lo hicieran de manera arbitraria y errática.

Si suponemos que el cumplimiento de los requisitos C-C-C es lo que distingue una idea proveniente de una cualidad primaria de una idea proveniente de una cualidad secundaria, significa que hemos encontrado una distinción relevante en lo que hace a nuestras ideas: la sistematicidad en el cumplimiento de los requisitos C-C-C nos indica

que hay algo afuera que se comporta de acuerdo a dichas ideas. Pero, ¿puede sostenerse esta hipótesis?

Recordemos que en la aberración (C) se fallaba sistemáticamente en detectar el color de un objeto. Bennett nos dice que la cualidad secundaria que origina la idea de color puede no estar correlacionada con otras ideas: que no perciba el rojo localmente no genera aberraciones compensatorias en otros casos de percepción visual. Bien podríamos tener el espectro invertido en lo que hace a las modalidades de idea correspondientes a cualidades secundarias, pero no tiene sentido decir que hay un “espectro invertido” en lo que hace a las cualidades primarias. Podría estar percibiendo lo dulce como salado, lo rojo como verde y lo agudo como grave; pero sería imposible o altamente implausible percibir lo alto como bajo, lo cuadrado como redondo o lo sólido como blando.

Pero, ¿puede *realmente* pensarse en la posibilidad del espectro invertido en lo que hace a las ideas originadas a partir de cualidades secundarias?

Locke es ambiguo en lo que hace a la formulación de las cualidades secundarias. A veces, identifica dichas cualidades con las ideas generadas: así, la idea de amarillo es la cualidad secundaria del amarillo:

[Las cualidades secundarias son] ideas tales que no son más que otras relaciones con otras sustancias y, consideradas tan solo por sí mismas, no están realmente en el oro. (Locke, [1710] 1956, II, XXIII, § 37, pág. 299)

A veces, en cambio, la cualidad secundaria es una *potencia* de producir una determinada idea:

La amarillez no está realmente en el oro, pero es una potencia en el oro para producir esa idea en nosotros cuando se sitúa en la luz debida ([1710] 1956, II, XXIII, § 10, pág. 281)

Como observa Bennett (1988, pág. 143), la expresión “esa idea” es insatisfactoria, porque parece referirse ambiguamente tanto a la idea generada por la cualidad secundaria -la idea de amarillo- como a la cualidad secundaria misma -la potencia de producir la idea de amarillo-, con lo cual aun en la formulación de la cualidad

secundaria como potencia se está usando la noción de cualidad secundaria en el sentido anterior, e.d., identificando o confundiendo la cualidad secundaria con la idea.

Posiblemente, esta ambigüedad está presente en nuestro uso cotidiano de los términos que corresponden a cualidades secundarias. A veces decimos que algo “es amarillo”, y con eso queremos referirnos a ciertas propiedades del objeto que, en ciertas condiciones de luminosidad, provocarían en sujetos humanos la visión del color amarillo. A veces hablamos del amarillo como “la idea en nosotros”, como cuando decimos “yo lo veo amarillo”. En todo caso, en el primer sentido (la cualidad secundaria como potencia) estamos pensando en una *causa* de un estado mental determinado (que yo tenga la idea de amarillo se explicaría aduciendo que hay algo que provoca esa idea); en el segundo sentido nos referimos a la idea como *efecto* de esa causa. Locke parece hablar de cualidades como “la amarillez” simultáneamente en los dos sentidos, en tanto que causa y en tanto que efecto, y por eso resulta ambiguo en su formulación. No es de extrañar este uso, atendiendo a la *disunidad* de la idea que señalamos en § 1.1. Si mantenemos ambos sentidos separados, entonces, tal vez no haya problemas en su tratamiento. Sin embargo, surge un nuevo problema: estamos aceptando una tesis causal en lo que respecta a cualidades secundarias. Esto implica que dichas cualidades sólo pueden instanciarse si se da el componente causal correspondiente. Y por ello mismo, resultaría contradictorio que se instancie, por ejemplo, una idea de color rojo sin su correspondiente componente causal. Esto lleva a afirmar que la posibilidad expresada en el caso de (C), quien poseía una aberración perceptiva que lo llevaba a afirmar sistemáticamente que un objeto era “rojo” cuando en verdad era “verde”, sólo podía darse si se modificaban las causas que producían la idea de verde en lugar de la idea de rojo. Pero en tanto las causas se mantengan idénticas, no hay espacio efectual para que en (C) se genere la idea de verde cuando se instancia el elemento cualitativo causal para producir la idea de rojo. El componente causal presente en la idea derivada de cualidades secundarias nos juega una mala pasada: ya no podemos establecer tan nítidamente la distinción que manteníamos entre cualidades primarias y secundarias, pues es necesario admitir que la huella causal está presente tanto en ideas provenientes de cualidades primarias como de secundarias. Locke, sin embargo, oscurece la distinción en otros dos sentidos:

- 1) Afirmando que las cualidades secundarias “no están en los objetos” (Locke, [1710] 1956, II, VIII, XIV)
- 2) Afirmando que las cualidades secundarias pueden reducirse a las primarias (Locke, [1710] 1956, II, VIII, § 17; II, VIII, § 23), mediante el avance en la investigación científica.

La afirmación 1) nos deja con una perplejidad: si las cualidades secundarias no están en los objetos, entonces su poder causal debe ser explicado por algo diferente a la propia existencia de los objetos. Presumiblemente, una parte de su poder causal es nuestra propia constitución sensorial. Ante esta última posibilidad, se puede entender la afirmación “no están en los objetos” de dos maneras diferentes: a) las cualidades secundarias “están en la mente”; b) las cualidades secundarias están en la interrelación entre el objeto y el sujeto percipiente. La opción a) convierte a la noción de “cualidad” en conceptualmente irrelevante: si la cualidad secundaria está en la mente, entonces es una idea o está a la par de una idea, y en ese caso no podría cumplir con el rol de intermediaria entre la mente y el mundo. Si tomamos la opción b), entonces, las cualidades secundarias *están* de algún modo y parcialmente en el objeto:

“Las cualidades secundarias sensibles, que dependen de [las primarias] no son otra cosa que las potencias que esas sustancias tienen de producir varias ideas en nosotros por nuestros sentidos; ideas que no están en las cosas mismas de otra manera que como cualquier cosa se halla en su causa” (Locke, [1710] 1956, II, XXIII, § 9, pág. 281)

Y también:

“Las cualidades secundarias (...) dependen (...) de las cualidades primarias de las partes minúsculas e insensibles [de los objetos]” (Locke, [1710] 1956, IV, III, § 11, pág. 543)

Se afirma, entonces la dependencia de las cualidades secundarias con respecto a las cualidades primarias: si veo rojo, es porque hay algo extenso, de determinada figura y tamaño que es causalmente suficiente para generar la idea de rojo, aunque soy incapaz de descubrir con claridad esa extensión, esa figura y ese tamaño. Para toda cualidad secundaria, hay al menos una cualidad primaria (aunque tal vez imposible de percibir) subyacente. Esto nos llevaría a decir que las cualidades secundarias son sólo cualidades primarias desconocidas, presentadas bajo una forma engañosa o poco

informativa. De allí la afirmación 2): el avance en la investigación científica nos iría mostrando que las cualidades secundarias sólo son cualidades primarias de corpúsculos, con lo cual el conocimiento microscópico cada vez más fino serviría para reemplazar cada cualidad secundaria por un conjunto de cualidades primarias. Una vez más, si esta es la interpretación que debemos dar al término “cualidad secundaria” (un conjunto de cualidades primarias imperfectamente conocidas), entonces no queda espacio para una “inversión de espectros”, a menos que el componente causal de la cualidad secundaria presente en la primaria no sea completo. En síntesis:

- Si Locke afirma que las cualidades secundarias son potencias de los objetos para producir sensaciones, y
- Si además afirma que las cualidades secundarias son cualidades primarias defectuosamente conocidas; entonces:
- Toda idea derivada de una cualidad secundaria deriva, transitivamente, de una cualidad primaria; por ello:
- Todas las ideas son derivadas, con mayor o menor informatividad, a partir de cualidades primarias y por lo tanto pueden reducirse a ellas.
- Si la inversión de espectros no era posible en lo que hacía a las ideas derivadas de cualidades primarias, dicha inversión de espectros tampoco parece posible en lo que respecta a las cualidades secundarias y, por lo tanto, *ambos* tipos de cualidades deben cumplir con los requisitos C-C-C.

Esta última consecuencia nos deja sin una base clara para distinguir las cualidades primarias de las secundarias. Sin embargo, caer en la identidad entre tipos de cualidades no hace justicia a la tensión que muestra Locke en sus usos de las expresiones referidas a cualidades secundarias: en pasajes sucesivos dice que dichas cualidades “no están en los objetos” y que “están causalmente en los objetos”. ¿Qué puede significar “estar causalmente en X”, aun sin estar en X?

Bennett expresa el aspecto distintivo de una cualidad secundaria a partir de lo que llama la “tesis analítica” de la causación de ideas a partir de cualidades secundarias:

Si x estuviese en la relación R con un humano normal, el humano tendría una idea sensorial de un tipo tal y cual (Bennett, 1988, pág. 128).

Decir ‘una manzana es verde’ equivale a decir ‘si la manzana estuviera iluminada con luz diáfana, y un humano tiene los ojos abiertos en dirección a ella, entonces vería la manzana de color verde’. A partir de esta tesis se le pueden conferir los siguientes rasgos a las afirmaciones del tipo “x es verde” (Bennett, 1988, pág. 139):

- 1) *Disposicionalidad*: “x es verde” equivale a un condicional contrafáctico de la forma: “Si x no hubiese estado iluminado por una luz diáfana, y un sujeto humano no lo hubiera estado mirando, entonces no habría una idea de verde”
- 2) *Carácter relacional*: “x es verde” no es una afirmación acerca de un objeto, sino acerca de sujetos. “x es verde” puede hacerse falsa si cambian las características perceptuales del sujeto, aun cuando no cambiase la estructura del objeto.
- 3) *Carácter mental*: “x es verde” significa algo acerca de *darse* una idea (no acerca de un objeto)

No todas las ideas derivadas de cualidades secundarias cumplen simultáneamente con estos requisitos, aunque necesariamente habrán de cumplir con al menos uno de ellos. Pero estos requisitos sólo nos hablan de un solo aspecto (la idea en el sujeto; es decir: el efecto), pero no acerca de su *causa*. Para poder enlazar el aspecto causal a la cualidad secundaria, necesitamos que la causa no sea ella misma una cualidad secundaria. El único candidato para tal aspecto causal es la cualidad primaria. La tesis causal que propone Bennett es, entonces:

Para alguna cualidad primaria \emptyset no disposicional: x es \emptyset y es una ley causal que si un ser humano se enfrenta a la luz del día con algo que es \emptyset , el humano tiene una idea G. (1988, pág. 140)

En rigor, Locke deja abierta otra posibilidad: quizás las causas mediatas de las ideas derivadas de cualidades secundarias no sean las cualidades primarias, sino *algún tipo de cualidad desconocida*:

“Las cualidades secundarias... dependen... de las cualidades primarias de las partes minúsculas e insensibles; o, si no de ellas, *de algo aún más remoto de nuestra comprensión*” (Locke, [1710] 1956, IV III § 11, pág. 543, cursivas añadidas).

Especulemos brevemente acerca de esta otra posibilidad. Ese “algo aún más remoto de nuestra comprensión” podría ser un tipo de cualidad no fenoménicamente

accesible, o una intervención divina (En el apartado 1.6 analizaremos esta última posibilidad). Mediante la aceptación de esta “cualidad inaccesible” podemos seguir manteniendo una tesis causal, pero estaríamos entrando en un juego peligroso. Esta afirmación abre la posibilidad de una brecha explicativa, precisamente la que se pretende cubrir con los conceptos de “cualidad primaria” y “cualidad secundaria”. La función de estos conceptos es la de explicar la relación entre las ideas y los cuerpos, entre las esencias nominales y las esencias reales. La cualidad secundaria sería, en este esquema hipotético, un tipo híbrido entre las ideas y otra cualidad desconocida. Esa otra “cualidad desconocida” es la brecha del puente que no podemos tender entre la realidad exterior y las ideas, puente que deberían tender las nociones de ‘cualidad primaria’ y ‘cualidad secundaria’.

Volvamos a la tesis causal manteniendo la suposición de que Locke acepta el enlace causal entre cualidades secundarias y cualidades primarias (y no un tipo de causalidad entre cualidades desconocidas y multiplicables) ¿Cuál es el justificativo para aceptar esta dependencia causal? Locke pareció entender que el avance de la ciencia no haría sino mostrar que todas las cualidades secundarias son explicables en términos de cualidades primarias: la ciencia efectuaría un análisis reductivo y explicaría el conjunto de propiedades de los cuerpos como un conjunto único de propiedades. Este ideal de unidad y reductividad está en la base de los intentos explicativos de Locke y Descartes. Bennett (1988, pág. 138) atribuye este optimismo al éxito de la física newtoniana, aunque Bennett aclara: “*Prima facie*, esta confianza causa perplejidad. Ciertamente no la garantizaba el estado de la fisiología del siglo diecisiete”. Probablemente no la garantizaba, pero la física newtoniana había logrado algo que ninguna teoría científica había conseguido hasta entonces: pudo explicar una infinidad de fenómenos de la naturaleza partiendo de leyes simples y elegantes, y manteniendo la suposición de la existencia de *corpúsculos*. En definitiva, la tesis causal parece apoyarse fuertemente en esta noción corpuscular y la posibilidad de reducción de cualquier fenómeno mental a interacciones mecánicas entre corpúsculos: la idea de rojo no es más que el producto de la percepción de una superficie compuesta por cuerpos pequeños cuya extensión, figura, número y reposo o movimiento no pueden percibirse a causa de tal pequeñez. La constitución sensorial del ser percipiente resuelve la imposibilidad de percibir las

cualidades primarias de la superficie “creando” el color a partir de los matices de luz. En rigor no hay una “creación” (el uso de este término nos comprometería a decir que las ideas producidas a partir de cualidades secundarias son, de algún modo, ficticias); la idea de color no es más que el producto necesario de una imperfección sensorial; es *lo que conocería alguien que no puede afinar la vista lo suficiente*. La idea de color es el modo como la mente llena un espacio epistémico en blanco. Por otra parte, la física corpuscular se erige como la candidata para explicar la regularidad de las percepciones: si las cualidades secundarias son relacionales, las potencias que producen en nosotros determinadas ideas deben seguir patrones que tengan conexiones legales y no meras arbitrariedades. Estas presunciones legales le permiten a Locke especular acerca de una conexión entre la física del mundo, la fisiología de los individuos (el aspecto relacional que se rige por las mismas leyes físicas que el mundo) y las ideas. Quizás, si prescindimos de la tentación de invocar cualidades “aún más remotas a nuestra comprensión”, podamos explicar, mediante patrones legales, la relación entre el aspecto fenoménico y el mundo externo.

1.2.2. Tendiendo el puente.

Intentamos encontrar el punto de conexión entre las ideas y el mundo, pues una vez encontrada la conexión causal de las ideas en el mundo, podremos introducir una cuña para explicar a las ideas como efecto de ciertos acaecimientos del mundo. De aquí nos hemos visto llevados a las nociones de *cualidad primaria* y *cualidad secundaria*. Hemos tratado de dar cuenta del aspecto causal del contenido objetivo (representacional) de las ideas y establecimos tres requisitos (los requisitos C-C-C) que satisfacen las ideas provenientes de cualidades primarias y que no satisfacen las que provienen de cualidades secundarias, lo que nos suministraría una base para concluir que las ideas provenientes de cualidades primarias contienen un indicio seguro acerca de su origen externo. Supusimos, junto con Bennett, dos sujetos que poseen aberraciones perceptivas y concluimos que las aberraciones con respecto a las cualidades primarias provocan, a su vez, aberraciones en todo el campo perceptivo, aunque esto no tiene por qué ocurrir con las cualidades secundarias. Los intentos de Locke para explicarnos

en qué consiste una cualidad secundaria (a diferencia de una cualidad primaria y de una idea) son insatisfactorios y problemáticos, pero construyen un marco conceptual en el que se da lugar al aspecto relacional - disposicional de las ideas surgidas a partir de cualidades secundarias. Esta disposicionalidad que involucra leyes permite inferir la existencia de patrones perceptivos y no meras arbitrariedades creadas espontáneamente por una conciencia. Sin embargo, Locke supone que las cualidades secundarias pueden reducirse a conjuntos de cualidades primarias, con lo cual la distinción entre cualidades se ve debilitada.

Podemos establecer relaciones entre patrones perceptivos. La concurrencia y persistencia de determinados tipos de cualidades (primarias y secundarias) nos permite establecer conjuntos de patrones, y a estos persistentes conjuntos Locke los denomina *esencias nominales* (Locke, [1710] 1956, III VI § 2, pág. 430; § 7-8, pág. 434). La esencia nominal es una idea abstracta formada a partir de ideas simples de conjuntos de cualidades. Los nombres de géneros naturales se refieren a su esencia nominal y no sólo a las ideas por ella causadas. Por ejemplo, la palabra “agua” no se refiere única ni paradigmáticamente a un conjunto de ideas asociadas a las vivencias que concurren en el acto de notar la idea de agua (la liquidez o solidez, la capacidad de mojar o de quitar la sed, etcétera), sino al conjunto de cualidades que provocan esos notares (aun cuando ese conjunto de cualidades sólo sea una referencia secundaria del término “agua”, dado que, como corresponde al internismo semántico, todo término tiene su referencia primaria en una idea). La percepción nos suministra el cumplimiento de los requisitos C-C-C: dada la concurrencia de cualidades coherentes, continuas y capaces de permitir inferencias perceptivas (situaciones perceptivas diferentes en las que el agua posee un conjunto de propiedades concurrentes¹⁵), parece natural agrupar las ideas derivadas de dichos conjuntos para darles el estatus de una sola idea compleja que merece llevar un nombre común (de ahí la suposición

¹⁵Para ser más exactos, el conjunto de propiedades concurrentes tiene la estructura de una disyunción inclusiva, pues la esencia nominal puede instanciarse aun cuando no siempre se instancien las mismas propiedades, aunque debe existir cierto conjunto mínimo de propiedades que debería instanciarse para decir que cierta esencia nominal se ha instanciado. Esto, por supuesto, es sumamente problemático, pues se puede pensar en casos en los que se presentan dos objetos con propiedades perceptivas totalmente diferentes, aunque les reservásemos a ambos tipos de objetos el mismo nombre de género natural. Es también conceptual y empíricamente posible la inversa: dos tipos de objetos que posean idénticas propiedades, pero que sean clasificados bajo esencias nominales diferentes. Cf. Locke, [1710] 1956, III VI § 31.

de que el nombre “agua” indique la esencia de algo, precisamente aquello que permite la concurrencia de ciertas ideas). Por ello, la esencia nominal es producto de una inferencia, y aunque se identifica con una idea abstracta, apunta directamente a las cualidades concurrentes que son causa de las regularidades perceptivas de un determinado aspecto de la naturaleza. La palabra “tigre” se refiere, primariamente, a las ideas mismas, pero se conecta estrechamente, de modo secundario, con un conjunto de cualidades que provocan típicamente ciertas ideas (determinado pelaje, color, cantidad de patas, etc.). La causa de estas regularidades entre patrones perceptivos es, precisamente, la esencia real. Pero esta es desconocida. Surge entonces el problema de la *legitimidad* de esta referencia secundaria (la esencia real), dado que sólo tenemos acceso a las ideas, y el lenguaje, según las palabras de Locke¹⁶, sólo puede referirse primariamente a ellas.

Afirmar de un objeto que cumple con determinada esencia nominal equivale a decir que dicho objeto tiene determinada estructura interna (García Carpintero, 1996, pág. 119). Pero la esencia nominal no habla de la estructura del objeto, sino de las cualidades presentes en la percepción: para poder afirmar algo acerca del objeto se requiere del concepto de *esencia real*. La esencia real es la referencia secundaria de los términos de género natural; esto es: es la afirmación de que las cualidades concurren en un objeto externo el cual, en principio, pudiera tener no sólo muchas más cualidades de las que se presentan en la esencia nominal, sino incluso podría poseer propiedades totalmente desconcertantes y diversas¹⁷. Locke, sin embargo, no parece

¹⁶ “Los nombres comunes de las sustancias, al igual que los demás términos generales, significan clases: lo que no es sino el ser signos de ideas complejas tales que en ellas se conformen o puedan conformarse varias sustancias particulares, en virtud de lo cual son capaces de quedar comprendidas en una significación común, y de ser significadas por un nombre” (Locke, [1710] 1956, III VI § 1, pág. 430) “La esencia de cada clase es la idea abstracta. La medida y el límite de cada clase o especie, por donde queda constituida en esa clase particular y distinguida de las demás, es eso que llamamos su esencia, que no es sino la idea abstracta a la cual va anejo el nombre, de manera que todo cuanto esté contenido en esa idea es lo esencial a esa clase” (§ 2, pág. 430) “Como las clasificaciones que nosotros hacemos de las cosas, o la determinación de especies, es con objeto de darles un nombre y de comprenderlas bajo un término general, no alcanzo a ver cómo pueda decirse con propiedad que la naturaleza establece los linderos de las especies de cosas; o si así fuera, nuestros linderos de las especies no se conforman exactamente a los de la naturaleza” (§ 30, pág. 450)

¹⁷ “Si quisiera hablar con alguien acerca de una clase de aves que últimamente vi en el parque de Saint James, de tres o cuatro pies de altura, cubierto de algo entre plumas y pelo, de un color café oscuro, sin plumas y sin alas, pero en lugar de ellas, dos o tres pequeñas ramas apuntando hacia abajo como renuevos de retama, con largas y grandes piernas, con patas de sólo tres garras, y sin cola, es necesario que haga yo esta descripción a fin de darme a entender por otras personas. Pero cuando se me ha dicho

afirmar que tales cualidades de la esencia real sean por principio incognoscibles, sino que el avance del conocimiento científico irá revelando características cada vez más finas de la composición corpuscular de los objetos¹⁸. Que un análisis tal sea alguna vez exhaustivo es apenas un *desiderátum*: jamás se puede estar seguro de haber conocido todas las cualidades de una esencia real.

La esencia real, aunque desconocida y solamente inferida, permite a Locke establecer una diferencia relevante en lo que respecta a conjuntos de ideas originadas en similares conjuntos de cualidades, pero a las cuales les corresponde una esencia diferente. Supongamos que llamamos “gato” tanto a los felinos como a cierto tipo de robots extraterrestres que tienen todas las cualidades que le adjudicamos a la esencia nominal “gato”. Puede que ningún análisis actual exhaustivo revele que se trata de dos tipos de seres diferentes: en ese caso, agruparemos a ambas “especies” bajo el mismo conjunto de cualidades y nos referiremos a ellas mediante una misma esencia nominal, aun cuando la esencia real sea diferente. Por ello, nunca tendremos la garantía de que hemos encontrado la esencia real de una determinada porción de la naturaleza. Ahora bien, podría ocurrir que a medida que se avance en la investigación científica, encontremos nuevas cualidades que vuelvan obsoleta la clasificación mediante esencias nominales. Para seguir con el ejemplo de “gato” y su contraparte robot, supongamos que al día siguiente se descubre la naturaleza robótica en algunos ejemplares de los que hasta hoy considerábamos “gato”: esto nos llevaría a afirmar dos esencias reales allí donde, hasta hoy, había sólo una. Seguiríamos llamando “gato” al felino, y (por ejemplo) “Robato” al gato robot, con lo cual habríamos trazado una nueva clasificación en la cual habrá dos esencias nominales. Pero esto no nos garantiza que a esas esencias nominales les corresponda sendas esencias reales: pasado mañana podría aparecer un nuevo descubrimiento, en el cual algunos de aquellos seres a los

que el nombre de ese animal es *casuario*, entonces ya puedo emplear esa palabra para significar en conversación toda la idea compleja aludida en aquella descripción, aunque por esa palabra, que ahora se ha convertido para mí en un nombre específico, no sé nada más acerca de la esencia real o constitución de esa clase de animales, de lo que abía antes de conocer ese nombre” (§ 34, pág. 453)

¹⁸ Aunque el conocimiento puede hacerse más fino, Locke es escéptico con respecto al conocimiento último y exhaustivo de las propiedades de la esencia real. Los mecanismos submicroscópicos están vedados al conocimiento humano, aunque se discute si Locke quiso decir que tal posibilidad está clausurada de hecho o si se trata de una limitación a priori del conocimiento humano. Cf. Yost (1951): “Speaking generally, Locke said that the observable clues are so scanty that there is no hope of making good guesses about *any* kind of specific sub-microscopic mechanisms” (pág. 126)

que llamamos “robotos” no eran robots sino seres vivientes extraterrestres. Y, paralelamente, quizás descubramos que dentro de los “gatos” haya algunos seres que poseen cualidades peculiares y distintivas, por las cuales debemos darle un nombre acorde con tales cualidades. El descubrimiento de nuevas cualidades esenciales no tiene por qué detenerse, y podría ocurrir que cada ejemplar que se nos presenta contenga características tan peculiares que resulte imposible clasificarlo con el mismo criterio que otro ejemplar. Esta posibilidad de multiplicación clasificatoria muestra que la divergencia entre la esencia nominal y la esencia real abre una peligrosa brecha de conocimiento: todas nuestras clasificaciones generales podrían revelarse, en última instancia, como clasificaciones puramente idiosincráticas, en las cuales cada hecho perceptivo habría de recibir un nombre diferente, lo que sería el fin de las esencias nominales y el comienzo del reinado del nombre propio *à la* Funes el Memorioso de Jorge Luis Borges (Borges, 1983). Locke no es totalmente escéptico en lo que hace al conocimiento de la esencia real. Parece necesario mantener la suposición de la existencia de una esencia real, para que el lenguaje tenga una referencia secundaria común lo suficientemente evidente como para que sirva a los fines de la comunicación¹⁹, aunque hay una restricción importante: si bien, según su punto de vista, podemos conocer propiedades de la esencia real (en tanto nuestras propias ideas contienen alguna referencia causal a estas propiedades, aunque las propiedades en sí son desconocidas), nunca podemos decir: “estas son todas las propiedades”. Pero precisamente esta imposibilidad hace que nuestras esencias nominales sean puramente tentativas y fundamentalmente inadecuadas.

Una opción que Locke no descarta es la de que *no* exista algo así como una esencia real²⁰. Esto no necesariamente debería embarcarnos en un lenguaje puramente de

¹⁹ “Los hombres, observando ciertas cualidades que siempre existen unidas, han copiado en eso a la naturaleza, y de unas ideas así reunidas han formado sus ideas complejas de substancias. Porque, si es cierto que los hombres pueden formar las ideas complejas que les vengan en gana, y pueden darles los nombres que se les antojen, sin embargo, si quieren darse a entender al hablar de las cosas que realmente existen, tienen, en algún grado, que conformar sus ideas a las cosas de las que pretenden hablar, porque, de lo contrario, el lenguaje de los hombres sería como el de Babel; y puesto que las palabras empleadas por cada quien sólo serían inteligibles para quien las empleara, ya no servirían para la conversación y para los negocios ordinarios de la vida, si las ideas significadas por ellas no respondieran de alguna manera a las apariencias comunes y guardaran alguna conformidad con las substancias, según realmente existen” (Locke, [1710] 1956, III VI § 28, pág. 448)

²⁰Cf: García Carpintero, 1996, pág. 122; Locke, [1710] 1956, libro III, cap. VI, §§ 8-9; libro III cap. VI §§ 49-50; libro III cap. IX § XIII, libro III capítulo X, § 20.

nombres propios: podríamos seguir utilizando un lenguaje que se refiriera a esencias nominales, pero tales esencias nominales tendrían un componente arbitrario, o disposicional: al no existir esencias reales como referencia secundaria de los nombres de género natural, las clasificaciones en géneros y especies serían arbitrarias, o dependientes de la constitución de los individuos, o quizás una mezcla de ambos. Podríamos seguir hablando de esencias nominales como si se refirieran (secundariamente) a una esencia real, pero ello por propósitos puramente pragmáticos.

Ahora bien, dentro del internismo semántico e intencional, la situación epistémica debe invertirse con respecto a la trazada en el apartado anterior. Habíamos establecido que los requisitos C-C-C parecían proporcionarnos una base para distinguir un componente causal de cualidades primarias de uno de cualidades secundarias (aun cuando inmediatamente encontramos que esta distinción está debilitada por el propio Locke al afirmar que las cualidades secundarias son cualidades primarias imperfectamente conocidas). En rigor, las cualidades primarias implican un grado de inferencia mayor que las cualidades secundarias. Si bien Locke parecía otorgar primacía ontológica a las cualidades primarias, el camino de la evidencia parte de las cualidades secundarias y sólo a través de ellas (inferencialmente) se dirige a las cualidades primarias. Sólo desde allí podemos inferir la esencia real. En palabras de García Carpintero:

(...)[L]a aseveración central del realismo por representación lockeano es que el contenido de *todos* nuestros estados mentales es “inmanente”: conciernen directamente a características de nuestras vivencias. Notamos regularidades en estas vivencias, y en virtud de esas regularidades notadas en ellas las tomamos como signos naturales de características objetivas de estados de cosas: es decir, las suponemos nómicamente conectadas con un mundo objetivo, cuya naturaleza colegimos a partir de la estructura de nuestras vivencias. (1996, pág. 124)

En otras palabras: todo el componente causal que hemos venido buscando no es más que una inferencia a partir de nuestras propias vivencias. El sistema de Locke, aunque empeñado en encontrar un nexo causal entre el mundo exterior y las ideas, no puede escapar del solipsismo epistémico. Aunque Locke afirme que de hecho hay una conexión y se muestre (en virtud de su confianza en los resultados de la ciencia)

esperanzado en el progresivo conocimiento de las cualidades de los objetos, su sistema no permite establecer un puente claro entre el ámbito de la subjetividad y el ámbito de la objetividad. Toda vez que nuestras afirmaciones en las que se involucran términos de género natural no son más que afirmaciones concernientes a las esencias nominales, y toda vez que las esencias nominales sólo se refieren a conjuntos de cualidades que bien podrían ser arbitrarios (admitido el hecho de que las esencias nominales son construidas y que no necesariamente deben reflejar el “estado de la realidad”), la posibilidad del solipsismo epistémico acecha fuertemente dado que no existe un criterio independiente que permita decidir entre conjuntos de ideas. El realista por representación lockeano sigue encerrado en la habitación y toda estrategia para investigar lo que pasa afuera mediante la táctica de explorar lo que hay adentro se ve desbaratada por la irremediable heterogeneidad de la esfera de lo objetivo con respecto a la esfera de lo subjetivo, y por su propia aceptación de que es imposible salir de la habitación. Sus indicios no son verdaderos indicios; son inferencias de indicios o, mejor dicho, inferencias acerca de ciertas ideas las cuales, convenientemente interpretadas, podrían funcionar como indicios. Locke se ve en un callejón sin salida. O acepta un idealismo como el de Berkeley (y claudica en su intento de encontrar los nexos causales de la experiencia), o admite abiertamente un realismo gnoseológico.

1.2.3. Aberraciones y más aberraciones.

En el apartado anterior (§ 1.2.2) distinguimos entre la esencia nominal y la esencia real. Establecimos que en el sistema de Locke nuestro conocimiento está dirigido a las esencias nominales y que, en rigor, la esencia real es un supuesto imposible de conocer de modo exhaustivo y seguro. Estamos empeñados en encontrar el punto de contacto entre las ideas –los contenidos de la mente- y el mundo – habitado por cuerpos que a su vez están compuestos por corpúsculos-, y de ese modo llevar a cabo el intento de reducción de las ideas a acaecimientos del mundo. Sin embargo el mundo se aleja cada vez más. En este apartado pondremos en duda la afirmación de que cierto tipo de aberraciones sólo son posibles en lo que respecta a las ideas

provenientes de cualidades secundarias, pero resultan implausibles si las suponemos como provenientes de cualidades primarias. Con ello, los requisitos C-C-C pierden su sustento y se debilita aun más la distinción entre cualidades.

Hemos examinado el hecho de que las cualidades primarias cumplen con los requisitos C-C-C, y que las cualidades secundarias no lo hacen. El cumplimiento de tales requisitos nos serviría, simultáneamente, para distinguir entre cualidades y para admitir que existe algo externo. Pero si nos atenemos estrictamente a las afirmaciones internistas de Locke, sólo poseemos acceso a las ideas. Las cualidades, sean primarias o secundarias, sólo pueden ser inferidas a partir de la idea. Las cualidades primarias permiten el establecimiento del juicio acerca de hechos positivos del mundo: la sustancia extensa (los cuerpos) parece ser una inferencia legítima, dado el cumplimiento sistemático de los requisitos C-C-C. Las cualidades secundarias sólo permiten juicios relacionales, relativos a la constitución perceptual del ser humano. Pero esta distinción sólo puede sostenerse si se acepta a la sustancia extensa como un hecho fundamental de la realidad: dado que algunos conjuntos de ideas provenientes de cualidades primarias convergen unívocamente en un mismo hecho perceptivo (el requisito de *coherencia*); cada cualidad no puede modificarse sin afectar al resto (la modificación de una cualidad implica modificación en otras cualidades, lo que llamamos *continuidad*) y podemos establecer inferencias y predicciones fiables entre perceptos (el requisito de *derivabilidad perceptual*), presumimos que existe algo sustancial que provoca la experiencia derivada de esos conjuntos de cualidades. Pero quizás el cumplimiento de estos requisitos sólo signifique que nuestra constitución perceptiva capte la realidad de ese modo, y por lo tanto el verdadero “dato objetivo” se escape irremediabilmente. Si esto es así, será conveniente, en primer lugar, examinar la fuerza del argumento para aceptar los requisitos C-C-C como definatorios del carácter objetivo del conocimiento.

Volvamos al ejemplo de (T) analizado en § 1.2.1: nuestro hipotético caso de un sujeto que tiene una aberración sensorial gracias a la cual una cualidad primaria es sistemáticamente mal percibida. Según el requisito de *continuidad*, las interrelaciones existentes entre las distintas cualidades primarias exigen que, cuando una de ellas se vea modificada, el resto también se modifique. Pero es conceptualmente posible que

sólo se modifique *una idea* derivada de un tipo de cualidad ante ciertos estímulos específicos, sin que se modifique el resto. Imaginemos que (T) tiene una aberración sensorial por la cual percibe visualmente todo lo cúbico como esférico (y todo lo cuadrado como redondo). Si seguimos el principio de continuidad, toda su experiencia sensorial deberá modificarse masivamente ante la presencia de cualquier objeto cúbico: necesitará de aberraciones compensatorias para justificar casos en los que por otro sentido (el tacto) podría verificar los vértices de un cubo que no podría verificar con la vista. Concedamos también una aberración compensatoria: ya no es sólo la vista la que participa de la aberración; el tacto también entra en el juego y también se ve influido por una extraña incapacidad selectiva para percibir cubos. Quedan más desafíos: si se ve un objeto cúbico como si fuera esférico, necesariamente no se verán los vértices del cubo y parece forzoso que, en su lugar, deba verse alguna otra cosa (figura 2) :

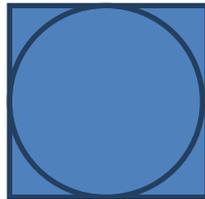


Figura 2.

Es evidente que quien no pueda percibir vértices en los cubos, pero pueda percibir todo lo demás (incluidos vértices en figuras que no fueran cúbicas) deberá compensar de algún modo esa imposibilidad. Alguien le podría indicar a (T): “Estoy señalando el espacio donde están los vértices que no puedes percibir. ¿qué es lo que percibes en este espacio?”. Aquí pueden ocurrir una de dos cosas: o bien (T) responde que no ve ese espacio (como si tuviera un escotoma selectivo que no le permitiera registrar los vértices de cubos), o bien (T) ve algo “en compensación” por esos vértices que no percibe. Lo mismo puede ocurrir con el tacto: quizás mientras está tocando un cubo, no puede “tener la idea” de vértice de rectángulo, a pesar de que nosotros vemos que está moviendo sus dedos en contacto con el vértice. Pero cualquiera sea su respuesta, puede mantener coherentemente que hay una esfera aun cuando produzcamos experimentos por otros medios para mostrarle que es un cubo: si posee exóticos

escotomas selectivos, tales escotomas pueden ser negativos, de modo que el sujeto no percibe la falta de aquello que no está percibiendo. En otras palabras, no percibirá que “le faltan” porciones del campo perceptivo y en rigor tampoco puede entender qué es lo que le queremos señalar. Es como si alguien se esforzara en hacernos creer que una esfera es en realidad un cubo, mediante el método de pasar la mano por algunos sectores de la esfera. Si percibe visualmente algo “en compensación” por los vértices que no percibe, aquí también puede darse el caso de que no note la aberración en su estructura perceptiva: puede ser que su campo visual se deforme de modo que los vértices sean simplemente omitidos, y en el lugar del vértice se podría percibir el fondo del paisaje en el que está el objeto²¹. Supongamos, para dar más dramatismo a nuestro intento por “hacerle ver” el cubo, que le pedimos a (T) que haga rodar el cuerpo por una superficie plana. Quizás, ante este hecho, (T) sólo perciba que la esfera no rueda con lo cual establecería el principio de que hay algunas esferas que ruedan²² y otras que no ruedan, lo que no parece a simple vista ningún impedimento para seguir considerando como esferas a los cubos. Otras posibilidades que descartamos son: que T se niegue sistemáticamente a responder, que trate de evitar el contacto con cubos o que se comporte como si hubiera cubos a pesar de que dice no percibirlos. Supongamos que es un sujeto entusiasta y colaborador, y su conducta es totalmente normal: si se cumplen las condiciones conductuales, pero el sujeto insiste en que un objeto es esférico en vez de cúbico, no nos queda más que dictaminar que está haciendo algún tipo de juego de palabras. En rigor, ni él ni nosotros podemos saber si en verdad está percibiendo cubos o esferas. *Todos estos casos nos llevan a concluir que las aberraciones compensatorias posiblemente eliminen el problema: estarán tan sistemáticamente ensambladas con el objeto percibido que el sujeto no tiene medios independientes para corroborar su aberración.*

Pero quizás con la aberración (T) se pretendía mostrar otro caso: (T) se da cuenta, por algún medio, de que no puede ver cubos: precisamente, de algún modo alcanza a percibir que no se cumple el requisito de coherencia en determinados tipos de percepciones con respecto a ideas derivadas de cualidades primarias. Locke examina

²¹No hace falta examinar esta posibilidad en lo que respecta al tacto, pues la experiencia táctil no necesita de un “fondo” que deba compensarse ante la falta de una modalidad perceptiva.

²²Mantengamos que (T) no tiene ninguna aberración en lo que hace a la percepción de una esfera.

este caso para comparar las propiedades de una idea derivada de cualidades primarias con una derivada de cualidades secundarias: "...si imaginamos que el calor, tal como está en nuestras manos, no es sino un cierto tipo un grado de movimiento en las partículas menudas de nuestro nervios o espíritus animales, podremos entender cómo es posible que la misma agua pueda producir al mismo tiempo la sensación de calor en una mano y la de frío en la otra. Lo cual, sin embargo, jamás acontece respecto a la forma, que *nunca produce la idea de un cuadrado en una mano, cuando ha producido la idea de un globo en la otra*" ([1710] 1956, II, VIII, § 21, itálicas añadidas) El ejemplo de Locke es engañoso: en realidad es conceptualmente posible (aunque empíricamente improbable) que una persona perciba algo redondo con una mano y algo cúbico con la otra. El caso no es más desafiante que el de quien percibiera rojo con un ojo y verde con el otro: será cuestión de encontrar algún tipo de experimento independiente que permita corroborar cuál ojo percibe "correctamente", o bien decidirlo por decreto emparejando la percepción de ambos ojos, haciéndoles percibir a ambos rojo o verde mediante un tratamiento correctivo. Estos casos involucran que el sujeto note la incoherencia presente en su percepción. Supongamos que (T) se da cuenta de su aberración porque, a diferencia de la vista, con el tacto sí percibe los vértices. Supongamos, también, que esta aberración le ha generado numerosos e insólitos problemas en su vida (por ejemplo, ha intentado jugar al bowling con cubos en lugar de pelotas) En este caso la violación del requisito de coherencia lo llevaría a sospechar que quizás algo anda mal en sus sistema perceptivo. Por ello, trataría de corregir sus errores de percepción, quizás pidiéndole a alguien que le indique si efectivamente algo es una pelota o un cubo. En definitiva, el problema de (T) es que cierto caso específico de instanciación de cualidades primarias es percibido de una manera por un sentido y de otra manera por otro sentido. Supongamos que existe una operación para corregir este problema. ¿Qué corregirá la operación?:

- a) Corregirá la vista para que sea coherente con el tacto; o bien:
- b) Corregirá el tacto para que sea coherente con la vista.

¿Qué puede preferir (T)? Probablemente dictamine que su problema está en la vista y no en el tacto. Puede que, en esta competencia sensorial, prefiera el dictamen de sus dedos antes que el de sus ojos: puede que crea que sus dedos le informan "la verdad",

mientras que sus ojos le “mienten”. Por ello, es posible que prefiera la operación (a). *Pero en todo caso, aun si se lo sometiera sin su conocimiento a la operación (b), (T) no se daría cuenta de que ha cambiado su modo de tocar antes que su modo de ver.* Mientras el requisito de coherencia se vea subjetivamente cumplido, no hay espacio para que el sujeto se dé cuenta de cuáles son los elementos que intervienen en la coherencia de sus notares²³. En el apartado 1.5. retomaremos el análisis y la posibilidad de las aberraciones.

Estas especulaciones pretenden mostrar que *el requisito de coherencia parece estar anclado en cierta tendencia espontánea a correlacionar de modo sistemático los diferentes tipos de ideas perceptivas*, más que a hechos objetivos. Que “objetivamente” algo sea cuadrado no tiene por qué provocar la modalidad perceptiva de la “cuadradez”: si se involucran coherentemente la totalidad de las modalidades perceptivas para crear de modo sistemático la idea de “redondez”, no se violenta el requisito de coherencia. Pero cuando Locke dice que las ideas derivadas de cualidades primarias son semejanzas (cf. § 19) parece estar anclando la coherencia en esta relación con lo semejante en la cosa real de la cual se deriva esa cualidad y que produce finalmente la idea. La pregunta que subsiste todavía es: ¿ha logrado salir Locke de la esfera subjetiva con la apelación a esta noción de semejanza? Es patente (y Bennett así lo enfatiza²⁴) que la noción de semejanza no puede tomarse en sentido literal, pues las ideas son por principio heterogéneas con respecto a las cosas (Analizaremos la noción de semejanza más abajo, en §1.5). En un sentido menos fuerte, la afirmación de semejanza se vuelve compatible con el realismo por representación: la realidad objetiva de la idea representa en algún sentido putativamente isomórfico a la realidad formal de su causa. Pero este sentido debilitado no cumple con el rol para el que se lo propone, porque debería poderse

²³Podríamos imaginar una tercera operación, en la cual al sujeto no se le corrigen los sentidos del tacto ni de la vista, sino que (mediante una complicada cirugía cerebral) se le “induce” la “sensación de coherencia” perceptiva, a pesar de que “en sí” las modalidades perceptivas son incoherentes. Pero en este caso tampoco queda espacio para seguir diciendo que las modalidades se mantienen incoherentes, porque de hecho son subjetivamente coherentes.

²⁴ “Puesto que las ideas no pueden asemejarse ni a los cuerpos ni a las cualidades de los cuerpos, esto o bien se descarta o bien se transforma. La única transformación plausible es en algo como lo siguiente: al explicar causalmente las ideas de las cualidades primarias, uno usa las mismas palabras al describir las causas como al describir los efectos” (Bennett, 1988, pág. 142), aunque cf. Jacovides [1999], para quien la semejanza debe entenderse de modo literal. Véase infra, § 1.5.

determinar por medios independientes que el carácter isomórfico se cumple. ¿Y de qué otro modo podría determinarse el isomorfismo entre representación y realidad si no es mediante otra representación? La alternativa que nos queda es que el isomorfismo es una hipótesis inferencial a partir de las ideas. El auténtico contenido representacional está en esas ideas, pero el isomorfismo es en sí mismo incorroborable desde lo representacional.

A Locke le parece que las ideas derivadas de cualidades primarias son en algún sentido más diáfanas que las derivadas de cualidades secundarias. Lo mismo afirmaba Descartes, al suponer que la noción de extensión es clara y distinta. Pero la claridad de estas nociones no nos resulta hoy evidente, y con el ejemplo de la aberración de (T) hemos intentado hacer plausible esa oscuridad. García Carpintero expone que las ideas simples provenientes de cualidades secundarias son lo más inmediato y, por lo tanto, lo más claro; pero en rigor se requiere de un argumento para aceptar esta claridad de las ideas como “cúbico” o “sólido”:

“Las ideas de propiedades primarias como # cúbico #²⁵, no son tan “diáfanas”. Las ilusiones perceptivas muestran que es posible que algo parezca un cubo a un ser humano normal y, sin embargo, no sea un cubo. Pese a ello, es parte fundamental de las ideas epistemológicas de Locke la creencia de que también las ideas de propiedades primarias son “diáfanas”, en el sentido de que se puede dar una explicación de las *condiciones normales* tal que si algo le parece cúbico a un ser cognoscitivamente equipado como un ser humano normal en circunstancias normales, es cúbico (y *ser cúbico* es parecerle cúbico a un ser humano normal en circunstancias normales) La determinación de qué son condiciones normales se haría de tal modo que quedarían excluidas las circunstancias en que se producen ilusiones perceptivas.”
(García Carpintero, 1996, págs. 123-4)

Tal como afirma García Carpintero, las propiedades primarias o, mejor, los conjuntos de cualidades primarias que conforman la estructura de un objeto requieren de la interposición de una cláusula en la que se presuponen las condiciones normales de los órganos del sujeto percipiente. Adicionalmente, se debería requerir de un criterio para decidir quién o quiénes están en esas condiciones normales. Los ejemplos que vimos más arriba, acerca de una “operación correctiva” de ciertas condiciones que se presumen anormales, pretendían otorgar plausibilidad a la tesis de que aun en

²⁵G. Carpintero estipula la convención de que todo lo que esté dentro de dos símbolos de número “#” corresponde al contenido de una idea.

condiciones groseramente anormales (aunque no manifiestas) es posible obtener una idea “errónea” de cualidades primarias siempre que se cumpla subjetivamente con el requisito de coherencia. Esto nos llevaría a cuestionar la semejanza entre la idea y la cualidad primaria correspondiente. Antes de ello, debemos evaluar el uso que hace Locke de la noción de ‘cuerpo’ (que contiene a la esencia real) de la cual se derivan las cualidades primarias, y la relación de reducción que existe entre cualidades e ideas.

1.3. Simetría de cualidades, asimetría de ideas y reducción.

Los cuerpos (aun cuando su esencia real sea desconocida) son los puntos de partida indiscutibles, y están presupuestos incluso en las afirmaciones con las que se pretende probar el carácter objetivo de las cualidades: Locke no duda en utilizar la noción de cuerpo en la carga de su prueba de que los distintos tipos de cualidades son, o bien cualidades reales de los cuerpos, o bien potencias de esos cuerpos para producir ciertas ideas (aun cuando no hubiera semejanza en ningún sentido potable del término). En el Ensayo Sobre el Entendimiento Humano, Locke nos dice:

El volumen, el número, la forma y el movimiento particulares de las partes del fuego o de la nieve están realmente en **esos** cuerpos, sean o no percibidos por los sentidos de alguien, y por eso pueden llamárseles cualidades reales, porque realmente existen **en esos cuerpos**. (Locke, [1710] 1956, II, VIII, § 17, pág. 116, negritas añadidas)

Cuando habla del fuego o de la nieve, parece estar hablando de la esencia nominal (y no puede entenderse de otra manera). Mientras se mantenga en la esencia nominal, está hablando de conjuntos de cualidades o, en rigor, de ideas producidas por conjuntos de cualidades. Pero inmediatamente utiliza la expresión “esos cuerpos”, como si no se estuviera refiriendo a tales esencias nominales, sino a la esencia real. Este es un salto injustificado: su argumentación sólo puede valer para las esencias nominales, y en rigor no puede salir de ellas. Sin embargo, mediante el uso ilegítimo del deíctico asociado a la noción de cuerpo, aparentemente Locke pretende saltar por encima de los conjuntos de cualidades para afirmar algo acerca de las esencias reales o cuerpos.

Más ejemplos. En II, VIII, § 22:

“...hace falta, para nuestra investigación, distinguir las cualidades primarias y reales de los cuerpos que **siempre están en ellos**” ([1710] 1956, II, VII, § 22, pág. 118 – 9, negritas añadidas);

Y en II, VIII, §22:

“...creo que pueden llamarse con propiedad *cualidades reales, originales o cualidades primarias*, porque **están en las cosas mismas**, se las perciba o no” ([1710] 1956, pág. 119 – 20, negritas añadidas)

En esta última cita se corrobora el salto presupuesto: Locke afirma algo con respecto a las esencias reales, lo que en principio es imposible (dado que el lenguaje con el que se realiza dicha afirmación sólo puede referirse a esencias nominales) y eso le hace girar en círculos: las cualidades primarias son reales porque están en las cosas mismas. Pero precisamente queremos saber qué significa “estar en las cosas mismas”. La única respuesta que podemos encontrar es: están en las cosas mismas porque son reales. Nuevamente, también, se salta de una afirmación acerca de conjuntos de cualidades o esencias nominales, a una afirmación acerca de cuerpos en los que tales conjuntos inhiere.

Podemos concluir a partir de estas inconsistencias que Locke puede afirmar sin contradicción que se cumplan los requisitos C-C-C, pero estos requisitos no bastan para otorgarle carácter objetivo a las ideas que cumplen dichos requisitos, *a menos que ya presupongamos la noción de cuerpo*. Pero si la existencia de los cuerpos ya está presupuesta de antemano, no necesitamos de dichos requisitos para diferenciar los tipos de cualidades, puesto que el propio concepto de cuerpo es el que hace todo el trabajo.

Por otra parte, el concepto de “cualidad secundaria” entendido como una cualidad primaria cuya magnitud, movimiento, volumen, figura o extensión no podemos percibir por la imperfección de nuestros sentidos, también se vuelve problemático²⁶.

²⁶ En el caso de los cuerpos muy pequeños, invisibles a los sentidos desnudos, debemos suponer una continuidad entre lo macroscópico y lo microscópico: Locke aplicará aquí una “regla de la probabilidad basada en la analogía”: “[L]as hipótesis acerca de fenómenos a nivel corpuscular deben concebir a las partículas insensibles y a sus propiedades como extensiones (con sólo diferencias de grado) de los cuerpos naturales visibles. La observación de estos últimos muestra [...] tal continuidad, tal conexión gradual entre ellos, que resulta inevitable pensar que lo más probable es que las leyes que rigen los fenómenos no observables sean muy parecidas (si no idénticas) a las que rigen los fenómenos que

En II, VIII, § 25 ([1710] 1956, págs. 120 -121) Locke nos dice que: “La sensación no descubre que en la producción de esas ideas [las derivadas de cualidades secundarias] contribuya el volumen, la forma o el movimiento de partes, y también porque **la razón no puede mostrar de qué modo los cuerpos por su volumen, su forma y su movimiento puedan producir en la mente las ideas de azul, amarillo, etc.**” (Negritas añadidas²⁷): Locke encuentra una heterogeneidad muy grande entre las ideas producidas por cualidades primarias y las ideas producidas por cualidades secundarias, y es llamativo que esta imposibilidad no sólo sea de los sentidos, sino también de la razón. Si apelamos a la confianza de Locke en una ciencia que descubra (aunque no exhaustivamente) las cualidades primarias causantes de nuestras ideas de cualidades secundarias, todo lo que podemos aspirar es a encontrar corpúsculos que posean cierta forma, extensión y movimiento, pero eso no nos explicará por qué lo vemos rojo o lo sentimos caliente, pero Locke parece consciente del hecho de que reducir un fenómeno a sus causas corpusculares no equivale a explicarlo en tanto que fenómeno de cierta cualidad definida. Presuntamente, el carácter fenoménico de la idea de rojo debería estar causado por la constitución fisiológica del sentido y no por la cualidad del cuerpo en el que se ve el color rojo. Pero el componente fenoménico de la idea de rojo o de calor no puede imputarse exclusivamente a la modalidad del sentido: si se afirmara esto último, las ideas fenoménicas no serían informativas acerca de hechos del mundo, ni podrían conectarse causalmente con él. De modo que el fenómeno cualitativo, producto de una cualidad secundaria, debe estar él mismo conectado con la cualidad secundaria (aun cuando no sea semejante a ella). Sin embargo, si las cualidades secundarias son solo cualidades primarias imperfectamente conocidas, podríamos esperar que una idea fenoménica producto de una cualidad secundaria pudiera descomponerse en una miríada de ideas producto de cualidades primarias desconocidas, aunque eventualmente cognoscibles. Por ejemplo, la realidad objetiva de la idea de una superficie roja podría descomponerse en un sinnúmero de realidades objetivas de ideas de partículas con determinada posición, figura, movimiento y

observamos cotidianamente”. (Lorenzo, 1996, pág. 122) (Cf. Locke, [1710] 1956, IV, XVI, § 12: “En las cosas que no pueden ser descubiertas por los sentidos, la analogía es la gran regla de la probabilidad”)

²⁷ Cf. con II, VIII, § 11 donde afirma taxativamente que “Las ahora cualidades secundarias de los cuerpos desaparecerían si pudiéramos descubrir las cualidades primarias de sus partes minúsculas”. Trataremos el rol de las hipótesis sobre inobservables más abajo, en § 1.7

solidez, ninguna de las cuales contiene algún aspecto de la realidad objetiva de lo rojo. Ahora bien, “rojo” o “caliente” suelen ser ejemplos clásicos de “ideas simples”. Desde un aspecto que involucra modalidades sensoriales, la realidad objetiva de las ideas de rojo o de caliente no pueden reducirse a otras realidades objetivas que involucren la misma modalidad sensorial: son primitivos fenoménicos. Para dar un ejemplo de esta noción de “primitivo fenoménico” tengamos presente la “idea de rojo”: la modalidad fenoménica del color no puede reducirse a una modalidad más básica. Esto significa que la “idea de rojo” es simple si, en el momento de efectuar una reducción, no encontramos otras ideas de la misma modalidad a la cual pudiera reducirse. En otras palabras: si fenoménicamente “rojo” no es equivalente a la conjunción fenoménica de otro fenómeno de color (por ejemplo, dado que según la explicación por pares oponentes, el rojo es un color unitario, por lo cual no puede descomponerse en otros fenómenos de similar modalidad, i.e., violeta y amarillo²⁸, entonces estamos ante un primitivo fenoménico. Pero quizás la “idea de rojo” podría descomponerse no en lo que respecta a su modalidad fenoménica (en tanto que color), sino en tanto tomamos a ese primitivo fenoménico como un derivado de otro tipo de modalidad. Si admitimos esta reducibilidad entre modalidades fenoménicas diversas, podemos aceptar que una idea simple como “rojo” pueda ser derivada de (reducida a) una idea de otra modalidad (por ejemplo, la posición, tamaño, figura y solidez de corpúsculos). En terminología cartesiana, es posible que la realidad objetiva de una idea se reduzca a las diversas realidades formales de otras. La reducción, sin embargo, no es *simétrica*: la conjunción de la realidad formal de las *ideas* de solidez, posición, tamaño y figura *no es idéntica* a la realidad objetiva de la *idea* de rojo, aun cuando pudiera haber simetría entre los tipos de *cualidades* a partir de las cuales se originan esas ideas. En síntesis:

- a) Aceptar que la “cualidad secundaria que provoca la realidad objetiva de la idea de rojo” sea reducible a “cualidades primarias de corpúsculos” nos permite aceptar que la conjunción de dichas cualidades primarias sea equivalente a la cualidad secundaria que instancia la idea de rojo. Por ello, podemos decir que la reducción entre *cualidades* es simétrica: una cualidad secundaria es el producto neto de una determinada conjunción de cualidades primarias, y una

²⁸ Cf. más adelante, en la segunda parte, Pautz (2006); Alex Byrne & Michael Tye (2006) y Foss (2009).

determinada conjunción de cualidades primarias produce o es una cualidad secundaria. La única diferencia en lo que hace a las cualidades secundarias es el componente subjetivo de la idea que provoca: la cualidad secundaria viene implicada por el contacto de los sentidos con las cualidades primarias. Pero “en sí”, una cualidad secundaria no es más que un conjunto abigarrado de cualidades primarias imperfectamente percibidas.

- b) Contrariamente, este principio de simetría no se cumple cuando se habla de las *ideas*. La realidad objetiva de la idea de rojo no es equivalente a la realidad formal de la idea de figura más la realidad formal de la idea de posición más la realidad formal de la idea de solidez. A su vez, la conjunción de dichas ideas no equivale al fenómeno de la rojez.

Si llamamos Q1 a las cualidades primarias y Q2 a las secundarias, podemos establecer la siguiente identidad:

$$Q1_1 + Q1_2 + Q1_3 + \dots + Q1_n = Q2$$

En cambio, si llamamos IQ1 a las ideas provenientes de cualidades primarias e IQ2 a las provenientes de cualidades secundarias, la identidad no puede establecerse, pues ninguna agregación de ideas IQ1 da como resultado una idea IQ2:

$$IQ1_1 + IQ1_2 + IQ1_3 + \dots + IQ1_n \neq IQ2$$

Esta asimetría en el plano de las ideas (pero no en el de las cualidades) es la que lleva a aceptar que una “idea de rojo” sea una idea simple: ninguna explicación exhaustiva a partir de ideas provenientes de cualidades primarias es *suficiente* para dar cuenta del aspecto cualitativo, fenoménico, de la idea de rojo.

1.4. El “micrango” y la relatividad de la reducción entre cualidades.

Las leyes de simetría expuestas en el párrafo anterior parecen indicarnos que para cada cualidad secundaria dada podemos encontrar un conjunto de cualidades primarias subyacente a las cuales puede reducirse, aunque esto no puede hacerse entre ideas. Mientras las cualidades podrían reducirse unas a otras, en las ideas hay un

componente que no puede reducirse sin pérdida de poder explicativo, y es precisamente la característica fenoménica lo que queda afuera.

Sin embargo, la aceptación de que las cualidades secundarias pueden reducirse a las primarias también tiene sus problemas, pues esto traería la indeseable consecuencia de que el estamento basal de la realidad no tiene ningún tipo de cualidades secundarias.

Veamos un ejemplo imaginario para exponer este problema de la reducibilidad de cualidades. De paso, el ejemplo nos sirve para ilustrar la asimetría entre ideas, aun cuando pudiera haber simetría entre cualidades. Supongamos que se inventara un “microscopio táctil”, cuyo efecto sería el de amplificar el fenómeno de la sensación táctil de una determinada superficie. Este dispositivo podría calzarse en una mano como un guante. Por medio de ese instrumento (para no violentar la etimología debiéramos llamarlo, quizás, “micrango”, reemplazando la raíz *-scopé* por *-agge*, ‘tocar’) tocaríamos con un dedo una pequeña superficie lisa de, por ejemplo, una mesa y obtendríamos información acerca de las pequeñas rugosidades e irregularidades de esa superficie. Gracias al micrango habremos reemplazado la sensación fenoménica de una modalidad táctil (la idea de liso) por una miríada de sensaciones de la misma modalidad, aunque con diferente –e incluso opuesta- cualidad (ideas de distintos grados de rugosidad). Una superficie que nos parecería lisa al tacto desnudo, se revelaría como porosa y rugosa tocada con el micrango. De ese modo, habremos reemplazado una cualidad táctil por su contraria gracias a la utilización de un instrumento que nos revela fenómenos microscópicos inaccesibles a la mano desnuda.

Supongamos ahora que una persona ciega utiliza el micrango para corroborar que a la idea de rojo le corresponden las ideas de una determinada posición, figura, tamaño y solidez de los corpúsculos que conforman la superficie roja: esta persona recibirá la información de todas las cualidades primarias de la superficie, pero la conjunción de esas cualidades primarias (aun cuando fuera resultado de una larga inferencia) no provocaría en él la idea de rojo²⁹. La persona ciega ha entendido perfectamente que el

²⁹ Nótese que con este ejemplo simplemente decimos que se trata de una persona ciega, sin declarar que tal persona tiene una especial capacidad de conocimiento como en el experimento mental de Mary (Jackson, 1982), la neurobióloga que lo sabe *todo* acerca del aspecto funcional de la visión y del cerebro.

rojo que no puede ver se reduce a la posición, figura, tamaño y solidez de los corpúsculos, pero esto no le da siquiera una pista de en qué consiste, fenoménicamente, el color rojo. Esto nos indica que, en lo que respecta al aspecto fenoménico de las ideas derivadas de cualidades secundarias, dicho aspecto fenoménico no parece derivable a partir de los componentes causales de las cualidades. Ninguna descomposición de cualidades secundarias en términos de cualidades primarias es suficiente para explicar de manera exhaustiva el aspecto fenoménico mismo. La persona ciega podría sacar las inferencias correctas acerca del color, pero eso no le informaría en lo más mínimo en qué consiste la cualidad fenoménica del color cuyas bases corpusculares está tocando.

Supongamos que el micrango tiene un nivel de detalle ajustable. Cuando el micrango está en un nivel de ajuste 1, tocamos una superficie lisa y descubrimos la (normalmente imperceptible) estructura rugosa de lo liso. Si lo ajustamos al nivel 2 (incrementando el poder de la “lupa táctil”), descubrimos que la disposición de cada uno de los corpúsculos tiene una superficie lisa (el micrango no está lo suficientemente calibrado todavía como para descubrir las rugosidades de esos corpúsculos). En el nivel 3, la textura lisa que habíamos descubierto en los corpúsculos del nivel 2 se revela como rugosa. En el nivel 4, los corpúsculos que provocaban la rugosidad del nivel 3 se revelan como lisos. ¿Es dable esperar que haya un nivel máximo de ajuste (más allá de las limitaciones del dispositivo) después del cual resulte imposible encontrar esta sucesión de liso – rugoso? ¿Estamos cometiendo un error de categorías si preguntamos por la cualidad táctil última de la materia o, -utilizando el lenguaje de Locke-, si preguntamos por las *potencias* últimas de los corpúsculos que pueden provocarnos una sensación táctil? La pregunta es puramente conceptual, porque va más allá de las limitaciones de cualquier dispositivo: aun si el micrango pudiera ajustarse al infinito, ¿habría un “punto fijo”, después del cual solo se transmite una única sensación (por ejemplo, liso) aun cuando se la siguiera ajustando indefinidamente? Y ese punto fijo sensorial, ¿estaría indicando el estadio último de la materia en el cual no hay más cualidades primarias sobre las cuales se pudieran percibir cualidades secundarias? En otras palabras:

- Las cualidades secundarias son potencias de producir una determinada idea. Por ejemplo, la (realidad objetiva de la) idea de *liso* es provocada por una determinada conjunción de cualidades primarias imperfectamente conocidas;
- A medida que se conocen más perfectamente las cualidades primarias intervinientes (i.e., a medida que aumentamos el nivel del micrango) a las que se reducen las cualidades secundarias, la idea provocada por cada nivel parece una idea diferente a la idea provocada por el nivel anterior (en un nivel es la idea de lo liso, en otro la idea de lo rugoso); de aquí:
- El conocimiento cada vez más exhaustivo de cualidades primarias no *elimina* a las cualidades secundarias en favor de cualidades primarias; mas bien hace que aparezcan *otras* ideas de cualidades secundarias: lo que era liso en el nivel 2 de ajuste, ahora es rugoso en el nivel 3.
- Aun conociendo exhaustivamente la esencia real, corpuscular, de una superficie, todavía podemos preguntar: “¿Y cuál es la idea derivada de una cualidad secundaria asociada a esas cualidades primarias últimas?”

Los últimos argumentos, el primero a partir de la asimetría entre ideas y el segundo a partir de la ajustabilidad de un instrumento hasta su grado máximo, pretenden abonar la conclusión de que las ideas derivadas a partir de cualidades secundarias se muestran como *irreducibles* a sus componentes causales de ideas primarias, y parecen *ineliminables* en sucesivos descubrimientos de cualidades primarias ocultas. Por lo tanto, los poderes causales de los cuerpos no parecen una condición suficiente para establecer la idea producida por cualidades secundarias en tanto que esa idea es informativa (aunque de un modo escaso o inadecuado) de cualidades primarias insensibles subyacentes. La afirmación anterior podría ser suscripta por Locke: no sólo se necesitan poderes causales para activar las ideas; también se requiere de modalidades sensoriales, y es por eso que las cualidades secundarias son esencialmente relacionales. Pero si seguimos el razonamiento de Locke (las cualidades secundarias son, en realidad, cualidades primarias imperfectamente conocidas), parece difícil aceptar que el conocimiento último de las cualidades primarias pudiera conllevar sensaciones cualitativas derivadas de cualidades secundarias. Supóngase que estamos observando el estadio corpuscular último de la materia –si tal cosa fuera

posible-: la posición, figura, la solidez y el movimiento de las partículas observadas serían las causas fenoménicas de las sensaciones de color o textura que obtendríamos en un nivel menos fino de observación (recuérdese que “idea de color” e “idea de textura” son derivadas de cualidades secundarias, las cuales sólo nos transmiten de modo imperfecto la estructura de cualidades primarias subyacentes); pero *en el nivel más básico posible no puede esperarse que haya color o textura: pues eso implicaría que hay otro nivel corpuscular más básico a partir del cual se derivan esas ideas de color o textura*. Aceptar un nivel básico nos lleva a aceptar la tesis de la segunda antinomia kantiana y, *a fortiori*, aceptar que si existe un nivel básico, ese nivel básico ha de ser necesariamente imperceptible, pues como seres sensoriales, tendremos que percibir alguna cualidad secundaria en ese nivel básico (y ya no puede haberla, porque una cualidad secundaria es, en realidad, una cualidad primaria subyacente mal percibida, y en el nivel básico ya no pueden haber cualidades primarias subyacentes. En cierto sentido, no puede haber una “percepción imperfecta” si se están observando los corpúsculos últimos de los que está hecha la realidad)

En síntesis (Figura 3):

<i>Nivel</i>	<i>Cualidad secundaria</i>	<i>Cualidad primaria subyacente</i>
1. Macro sensorial	Color de un cuerpo	Corpúsculos de determinada forma y posición, etc. iluminados.
2. Microscópico	Color de los corpúsculos del nivel 1	Corpúsculos más pequeños de determinada forma, posición, etc. iluminados.
3. Nivel mínimo de	No pueden poseer	Corpúsculos

la materia	color, pues no hay un nivel ulterior mediante el cual pueda analizarse.	mínimos.
------------	---	----------

Figura 3

La objeción que hemos interpuesto (que las ideas provenientes de cualidades secundarias son ineliminables en sucesivos afinamientos de conocimiento de cualidades primarias) está en la base de la objeciones de Geroge Berkeley. E.M. Adams (1948) dice al respecto:

Locke's basis for holding that our ideas of primary qualities resemble their corresponding qualities in the external object was that regardless of how we modify the object it still has these qualities. His example is that regardless of how many times we divide a grain of wheat each part will still have the capacities to produce in our consciousness the ideas of figure, size, motion, number and solidity. Berkeley was certainly correct in pointing out that the same is true with the secondary qualities and that it does not argue for objectivity in regard to either. For example, regardless of how we modify a physical object it will always have the capacity to produce a color sensation in us, unless, of course, it is made too small to be visible, but the same limitation applies to the primary qualities. (Adams, 1948, pág. 440)

La objeción de Berkeley menoscaba la posibilidad de una reducción de cualidades. Quizás la pregunta por el color de una partícula última no sea un error categorial. Pongamos el argumento de la siguiente forma:

1. Si las Q2 de un determinado nivel de la materia se reducen a Q1 de un nivel inferior, y
2. Si en cada nivel aparecen nuevas Q2 que serán ulteriormente reducidas al conocer las Q1 del nivel inferior, entonces o bien:

C1: En el nivel mínimo de la materia no pueden aparecer Q2 (pues no hay un estado ulterior al que pueda reducirse) y por lo tanto preguntar por el color de una partícula fundamental es un error categorial, o bien:

B1 (*Objeción de Berkeley*): En el nivel mínimo de la materia aparecerán *también* Q2 (y esto mostraría que las Q2 no son ulteriormente reducibles a Q1)³⁰

¿Cómo decidir entre C1 y B1?

De acuerdo con Locke y con el ideal científico del siglo XVII, el conocimiento total de la esencia real implicaría un tipo de explicación que no involucrara cualidades secundarias. Esta suposición se vio fortalecida durante el siglo XIX con la aparición de la teoría atómica de la materia y la teoría ondulatoria de la luz. Priest (1989, pág. 32) afirma que “with the help of these, it could be shown that the dispositions were really aggregate primary properties of the micro – structure of matter. It became clear that to say, *e.g.*, that fundamental matter, an atom, is coloured, is not just false, but a category mistake”. Priest, sin embargo, contrastará las teorías del siglo XIX con las afirmaciones acerca de la realidad de los fenómenos cuánticos. En particular, dará nueva vida a la distinción entre cualidades primarias y secundarias, interpretadas desde una óptica realista de la mecánica cuántica: existe una jerarquía de cualidades primarias y secundarias, relativas a diferentes niveles de la realidad. “The behaviour of macroscopic bodies and their properties is explained in terms of the (primary) properties of its microscopic (atomic) parts. The behaviour of these and their properties is, in turn, explained in terms of their quantum states and properties. Maybe the hierarchy goes further, though if it does, we haven't got there yet.⁷ Now traditional secondary properties are properties of matter at the macroscopic level; their perception in an observer is caused (in part) by a dispositional property to be understood in terms of structure at the microscopic level. Newtonian properties, on the other hand, are properties at the microscopic level (and so, derivatively, at the macroscopic level); their registering with an observer is caused (in part) by dispositional properties (or propensities) to be understood in terms of structure at the quantum level.” (Priest, 1989, pág. 36). Los elementos de la figura 3 podrían disponerse de la siguiente manera (Figura 4):

³⁰ Berkeley objetaría algo más, que para nuestro propósito puede ser omitido: no diría exactamente que en el nivel último de la materia hay cualidades secundarias, sino que tanto las cualidades primarias como las secundarias son ideas en la mente, y que por tanto la noción de “materia” es inerte en este razonamiento. Cf. Berkeley, (1710) 1985, XXXV, pág. 89

<i>Nivel</i>	<i>Propiedad característica</i>	<i>Producido por</i>	<i>Las disposiciones son:</i>
Macro	Secundaria	Disposición macroscópica más observador	Propiedades agregadas de estados primarios
Micro	Primaria (newtoniana)	Disposición microscópica más observador.	Propiedades vectoriales de estados cuánticos
Cuántico	Cuántica	?	?

Figura 4. (Extraída de Priest, 1989, pág.36)

Las diferencias entre la figura 3 y la figura 4 indican que las cualidades primarias y secundarias sólo son relativas a determinados niveles de la materia. Según lo esquematizado en la figura 4, no tendría sentido hablar de una cualidad primaria última e irreductible, puesto que las cualidades primarias no se resuelven mediante la apelación a la exhaustividad de una observación microscópica: es posible hallar nuevos tipos de cualidades primarias cuánticas subyacentes en las cualidades primarias microscópicas. Por eso, la propiedad cuántica de la figura 4 podría ser un nuevo tipo de cualidad primaria, la cual produciría nuevos tipos de ideas provenientes de cualidades secundarias (las celdas marcadas con signos de interrogación en la figura 4).

La lectura que realiza Priest acepta la posibilidad de una jerarquía infinita entre cualidades. Si en el esquema de Locke nos vimos llevados a aceptar la tesis de la segunda antinomia kantiana, con este esquema desembocamos en su antítesis. No parece implausible, ni contradice las esperanzas teóricas de Locke, que exista una jerarquía de cualidades en tanto estas sean entendidas como cualidades primarias. La propiedad cuántica de la figura 4 debe ser algún nuevo tipo de cualidad primaria (en la

que tal vez no se involucren los conceptos clásicos de “cualidad primaria”, esto es: número, figura, etc., sino algún otro todavía desconocido aunque eventualmente cognoscible mediante la investigación empírica) en tanto deseemos conservar el realismo con respecto a los fenómenos cuánticos.

Si tomamos la consecuencia de Priest, de que puede haber infinitos conjuntos de cualidades primarias que se reducen unas a las otras a distintos niveles, esto podría conllevar a que en cada micronivel aparezcan nuevos conjuntos de cualidades secundarias, de acuerdo a cuál es la cualidad primaria subyacente a ese nivel. Pero, como vimos más arriba: aunque existiera una simetría entre el plano de las cualidades (en el cual conjuntos de Q1 explican reductivamente conjuntos de Q2), esto no implica que haya tal simetría en el plano de las ideas. Conjuntos de ideas derivadas de cualidades primarias no son equivalentes a ideas derivadas de cualidades secundarias. Dado que, en lo que respecta a las cualidades secundarias, las modalidades del sujeto cognoscente intervienen activamente, es de presumir que esta asimetría se deba a que en el caso de las cualidades primarias, el sujeto sólo se limita a “recibir” el producto neto de las cualidades, sin agregados ni deformaciones; pero en lo que respecta a ideas provenientes de cualidades secundarias, todo lo que hay es precisamente una intervención por parte del sujeto más algún elemento causal anclado en las cualidades primarias. La apelación a la recepción puramente pasiva de una cualidad primaria pretende garantizar la objetividad de dicha recepción: el sujeto no “manipula” sus facultades sensoriales para generar un producto ideacional espurio. En cambio, la actividad es signo de “manipulación cualitativa”: el sujeto cognoscente deforma las cualidades primarias, las convierte en secundarias mediante el sentido y genera un producto que no posee semejanza con el objeto. La pasividad se convierte entonces en un signo de objetividad. He aquí, entonces, que a los ya de por sí insuficientes requisitos C-C-C debemos agregarle una cláusula: *el cumplimiento de los requisitos C-C-C nos suministrará un conocimiento de los cuerpos siempre que la res cogitans no participe activamente en la creación de las ideas que cumplan con esos requisitos*. En otras palabras, lo que deseamos es que las ideas provenientes de cualidades primarias se limiten a reflejar lo que ya está en el cuerpo, sin “agregados” de la propia mente o

del sistema sensorial. Esta división entre la pasividad-objetividad y actividad-subjetividad se volverá insostenible en poco tiempo, con el advenimiento del idealismo trascendental kantiano. Además, esta cláusula deberá refinarse para mostrar una diferencia sumamente implausible: que cuando se observa una cualidad primaria (por ejemplo, la forma de una mesa) ni los sentidos ni la propia mente “manipulan” de manera activa la percepción, mientras que en el caso de una cualidad secundaria la mente interviene de modo activo, y no existen criterios para determinar en qué casos hay actividad y en qué casos no la hay, excepto la determinación por decreto. Incluso, la sola admisión de este requisito de la pasividad es problemática, porque corre el riesgo de morderse la cola: las ideas derivadas de cualidades primarias reflejan objetivamente dichas cualidades debido a que se limitan a asemejarse pasivamente a ellas, y esto último es lo que les otorga el carácter de objetividad. ¿Cuál es el argumento para aceptar que la pasividad es un criterio de objetividad, o viceversa? Allí se requiere de un argumento ulterior. Locke a veces parece poner la carga de prueba en los órganos sensoriales: “Es claro que esas percepciones (i.e., ideas acompañadas de placer o de dolor) las producen en nosotros causas exteriores que afectan nuestros sentidos, porque aquellos que carecen de los órganos de cualquier sentido, nunca pueden tener producidas en sus mentes las ideas que pertenecen a ese sentido” (Locke, (1710) 1956, IV, XI, § 4, pág. 634). Pero, como señala Bennett (1988, pág. 92), esta argumentación contiene “una premisa acerca de órganos sensoriales (...) Pero los órganos sensoriales se encuentran entre las ‘cosas fuera de nosotros’ cuya realidad está en cuestión”. Nuevamente, la apelación a órganos sensoriales, lejos de disolver la petición de principio anterior, genera una nueva. La situación es paralela a la que habíamos encontrado en Descartes (*supra*, pág. 26): Los nervios, a diferencia de los espíritus animales, son emisarios del hecho perceptivo y reciben la estimulación inmediata de la exterioridad *de la que son parte*.

1.5 La semejanza entre cualidades primarias e ideas.

En el apartado anterior evaluamos la posibilidad de reducción entre cualidades, y concluimos que tal reducción es solo un supuesto derivado de la noción de cuerpo,

pero que en definitiva no hay argumentos sólidos para sostener tal reducción si se la lleva hasta el estado mínimo de la materia. Podría argüirse, con Berkeley, que aun los corpúsculos deberán tener cualidades secundarias, las cuales no serían ulteriormente reducibles. No hemos examinado, sin embargo, la idea de Locke de que las cualidades primarias semejan propiedades de los cuerpos, mientras que las cualidades secundarias no lo hacen. Si así fuera, sería posible afirmar que un corpúsculo mínimo tuviera color (tal como afirma la tesis B1), pero que ese color no fuera una propiedad del cuerpo, sino puramente del sistema sensorial del percipiente. Veamos, entonces, cómo puede aplicarse la noción de semejanza.

Las doctrinas del realismo por representación, como las de Descartes y Locke requieren de criterios para determinar la adecuación de una idea con un hecho externo. Hemos examinado que estos criterios (los requisitos C más la cláusula de pasividad) no cumplen con esa función, o bien porque los criterios están anclados en la subjetividad (como ocurría plausiblemente con los requisitos C), o bien porque involucran una petición de principio (tal como veíamos con la cláusula de pasividad, que a su vez involucraba una afirmación acerca de órganos sensoriales). Por lo tanto, la noción de *cualidad*, como auxiliar epistémico intermediario entre lo interno y lo externo no cumple con el objetivo de *informar* acerca de la esencia real, sino que en el mejor de los casos se limita a producir un efecto a partir de una causa puramente inferida. Podemos decir que de acuerdo al efecto, la causa puede ser evaluada a partir de esos mismos criterios que hemos puesto en discusión: si se trata de una idea derivada de una cualidad primaria, existe cierta semejanza entre la idea y una cualidad de la esencia real. Si se trata de una idea derivada de una cualidad secundaria, no hay semejanza debido a la intervención activa de la estructura cognoscitiva en la formación de una idea a partir de datos que escapan a los sentidos (como en el caso del color, cuya causa externa es una determinada disposición de corpúsculos y cuya causa interna es la actividad de la mente para producir una idea a partir de la imposibilidad de formarse una idea de la miríada de cualidades primarias insensibles de los corpúsculos). Sin embargo, hablando en rigor, si aceptamos que las cualidades secundarias son cualidades primarias imperfectamente conocidas; y si aceptamos que las ideas derivadas de cualidades primarias se asemejan a las cosas, entonces las ideas

derivadas de cualidades secundarias deberían *semejarse* en cierto sentido a las cosas. De otro modo, las ideas de color o de sonido no podrían ser informativas o, dicho más exactamente, no tendría sentido afirmar que en la exterioridad se encuentran las *potencias* de su producción (la idea es de algún modo un *reflejo* o un *producto* de esa potencia) y que esas potencias son en última instancia analizables como cualidades primarias.

Identificar las cualidades primarias de un vaso equivale a *tener* un vaso: las cualidades primarias no son potencias, son el objeto mismo. No se puede tener el conjunto de ideas (derivadas de cualidades primarias) que delimitan la forma y la estructura de un vaso sin tener la idea de vaso. Bennett es explícito en este punto:

La tesis analítica (cf. *supra*, pág 44) dice que el que una cosa tenga una cualidad secundaria dada es que tenga cierta potencia y sólo esto, prescindiendo de cualquier cuestión acerca de qué tipo de potencia, es inaplicable a las cualidades primarias de las cosas. Por ejemplo, podemos identificar el vaso e ignorar o estar en descuerdo acerca de sus cualidades secundarias; y así tenemos la noción de *el vaso* como un objeto que, entre otros hechos acerca de él, tiene ciertas “potencias” para afectarnos de ciertas formas que nos proporcionan la base para acreditarle color, sabor, etcétera” (...) “Es engañoso hablar de su potencia [*i.e.*, *del vaso*] para hacernos percibir que tiene cierta figura, tamaño, etcétera, pues esa forma de hablar sugiere que tenemos una noción de él –alguna forma de identificar y estudiar el vaso- independiente de y como preliminar para descubrir cuáles son sus cualidades primarias (Bennett, 1988, pág. 136)

De acuerdo a la afirmación de Bennett, con respecto a las cualidades primarias se sigue un camino directo entre éstas y la formación de la idea. Más arriba hablamos de un carácter pasivo de las ideas formadas a partir de cualidades primarias. Aquí tenemos una formulación un poco más explícita de dicha pasividad: no hay instancias identificadoras previas o independientes, con respecto a los vasos, que no sean las propias cualidades primarias. O, dicho de otro modo, las cualidades primarias otorgan la estructura a una idea, y una idea derivada de cualidades primarias no es más que (¿el reflejo de?) esa estructura. Toda idea adicional al conjunto de cualidades primarias de un objeto es producto de cualidades secundarias, las cuales son producidas activamente por la interrelación de la mente y la estructura insensible del objeto.

Podemos inferir que la noción de “*semejanza*” en lo que respecta a las ideas producidas por cualidades primarias consiste en una *identificación* entre la idea y la

cualidad, y esta identificación va más allá del puro isomorfismo representacional: la idea de vaso es *indiscernible* del conjunto de cualidades primarias que conforma esa idea. *Es indiscernible en tanto el conjunto de cualidades primarias no se presenta como algo diferente del objeto mismo*. Este conjunto puede conformar una esencia nominal. Las referencias secundarias de las esencias nominales son los tipos de cualidades (primarias y secundarias), pero mientras las cualidades secundarias están interferidas por la participación activa del sujeto, las cualidades primarias conforman por sí mismas ciertas estructuras nominales básicas, en las que las ideas conformadas por cualidades secundarias inhiere. En otras palabras, la idea de *rojo* es inseparable de la idea de *superficie*, en la que se da el color rojo. Con este ejemplo puede mostrarse que las ideas derivadas de cualidades secundarias parecen presuponer las ideas de cualidades primarias. Por eso, decir que *este vaso es verde* equivale a decir que en la idea compleja de *vaso*, derivada de un conjunto de cualidades primarias, está presente la idea de *verde*. “He sentido una rugosidad que no tenía extensión” o “he visto un verde que no era una superficie” parecen expresiones que contienen alguna clase de contradicción perceptual. Decimos que la esencia nominal de un objeto se identifica *estrictamente* con el conjunto de las cualidades primarias que conforman esa esencia nominal, pero se identifica de un modo *laxo* o *no estricto* con el conjunto de cualidades primarias *más* cualidades secundarias. Esto último significa que en la esencia nominal (EN) convive una disyunción inclusiva del conjunto de cualidades secundarias que podrían inherir en todos los posibles estados perceptivos:

$$EN = (Q1_1 \wedge Q1_2 \wedge \dots \wedge Q1_n) \wedge (Q2_1 \vee Q2_2 \vee \dots \vee Q2_n).^{31}$$

De esta manera, hablar *de un vaso* es hablar *necesariamente* acerca de un conjunto de cualidades primarias y *contingentemente* de un conjunto de cualidades secundarias. En una idea cualquiera de vaso, han de estar presentes, de manera necesaria, determinados caracteres primarios de un vaso, aunque puede haber una considerable variedad de caracteres secundarios.

Destacaremos que, a pesar de que –como sugerimos más arriba– una idea derivada de una cualidad secundaria parece inherir en una idea derivada de cualidad primaria, esto

³¹ La esencia real (ER) consistirá en el conjunto *exhaustivo* de cualidades primarias. Esto es: $ER = Q1_1 \wedge Q1_2 \wedge \dots \wedge Q1_n$, más la cláusula “Estas son todas las cualidades primarias subyacentes”.

no implica que la idea derivada de una cualidad secundaria pueda explicarse *completamente* en términos de ideas de cualidades primarias. Sin embargo, inherencia no implica *semejanza*. En todo caso, si existiera algún tipo de semejanza, sería en términos de las cualidades secundarias involucradas *más* la intervención del sistema perceptivo. Sin embargo, como veremos en el apartado 1.6, esto tampoco parece suficiente para afirmar semejanza entre una idea fenoménica y el sistema perceptivo más la cualidad primaria involucrada.

Averill (1982) enumera las diferencias características entre las ideas producidas por cualidades primarias y las producidas por cualidades secundarias. Estas diferencias han sido postuladas por Bennett (1988) como las marcas distintivas de las cualidades, y pueden resumirse en tres tesis:

- a) **La tesis de la dependencia** (*The Dependency Thesis*): las ideas producidas por cualidades secundarias dependen de un sistema perceptivo; no ocurre lo mismo con las ideas producidas por cualidades primarias.
- b) **La tesis de la ceguera** (*The Blindness Thesis*): es posible imaginar que una persona no puede distinguir entre dos ideas provenientes de cualidades secundarias (por ejemplo, entre rojo y blanco) y que por lo tanto sea “ciega” a esa distinción, pero no es posible imaginar esa ceguera en lo concerniente a ideas provenientes de cualidades primarias. (Por ejemplo, la distinción entre figura y movimiento)
- c) **La tesis disposicional** (*The Dispositional Thesis*): mientras las ideas provenientes de cualidades secundarias son completamente disposicionales (i.e., que X sea rojo equivale a decir que X es perceptible como ‘rojo’ para observadores que están en las condiciones apropiadas), las ideas provenientes de cualidades primarias no pueden analizarse exhaustivamente en términos de disposiciones perceptuales.

A partir de estas tres tesis debería ser posible afirmar que las ideas de cualidades primarias reflejan o asemejan una cualidad en el objeto, pero las ideas de cualidades secundarias no pueden hacerlo. Por ejemplo, dado que la noción de *figura triangular* se corresponde con una cualidad primaria, entonces es lícito concluir que la idea de

una figura triangular no depende de un sistema perceptivo (tesis de la dependencia), no es posible ser “ciego” a esa figura (tesis de la ceguera) y la figura subsiste más allá de que haya un sujeto en disposición de percibirla (tesis disposicional), de modo que una idea de figura triangular ha de reflejar una propiedad presente en la figura, esto es, la ‘triangularidad’. No ocurre lo mismo, por ejemplo, con la idea de ‘figura roja’, puesto que la ‘rojedad’ no es una propiedad de la figura, sino una relación que se establece entre un sistema perceptivo y ciertas potencias de los objetos.

Cuando en el apartado 1.3.2 hablamos de la posibilidad de aberraciones perceptivas, supusimos que una persona era “ciega” a cierta distinción de cualidades primarias (en el ejemplo dado, percibía lo cuadrado como redondo). La tesis de la ceguera nos dice que se puede ser ciego a ciertas cualidades secundarias, pero no a cualidades primarias, pues la ceguera en cualidades primarias se difundiría (*spreads*) a toda percepción debido al requisito-C de la continuidad: la modificación de una cualidad en un caso particular implica la modificación de todas las otras cualidades. En nuestro ejemplo mostramos cómo podría funcionar una ceguera determinada y cómo el requisito de coherencia podría mantenerse aun a pesar de tal ceguera. Averill muestra que la tesis de la ceguera no puede mantenerse *ni siquiera* como criterio distintivo entre cualidades primarias y secundarias, pues una aberración en una cualidad secundaria podría difundirse (*spread*) hacia otros casos inicialmente no contemplados. Supongamos que una persona tiene una aberración perceptiva que no le permite distinguir entre blanco y rojo:

A normal person could distinguish red objects from white objects by placing them under a strong green light, or by looking at them in sunlight after placing a heavily tinted piece of green glass over them, because under such conditions white objects appear green and red objects appear black. But the red-white color blind is unable to sort objects in this way, since he does not have any experiences that would allow him to distinguish red from white objects. Hence the green appearance of a white object and the black appearance of a red object, when seen under green light, look the same to the red-white blind. (...) [I]t follows that the red-white blind cannot distinguish green from black objects. The red-white blind is also green-black blind. (Averill, 1982, págs. 349 - 50)

Bajo condiciones de luz verde, un círculo blanco debería verse verde y un círculo rojo debería verse negro. Pero quien es ciego a la distinción rojo-blanco no podría,

manteniendo sus disposiciones, notar la distinción verde-negro, la cual no es más que la distinción rojo-blanco bajo otras condiciones luminosas. Pero, ¿qué ocurriría si cambiaran sus disposiciones? Quizás la tesis disposicional sea inadecuada:

Suppose that all of us became red-black blind so that ripe tomatoes and coal appear to have the same color. From the Dispositional Thesis it follows that

(1) *x* and *y* have the same color (taste, smell).

Means the same as

(2) If normal observers were to see (taste, smell) *x* and *y* under appropriate conditions for perceiving color (taste, smell), the *x* would appear to have the same color

(taste, smell) as *y* to those observers (Averill, 1982, pág. 352)

(2) no puede ser idéntico a (1), porque es posible encontrar casos en los que (1) sea falso y (2) sea verdadero: “Professional wine tasters can taste differences between wines that normal people cannot taste. Presumably this is because such wines have different tastes, although they taste the same to normal observers” (1982, pág. 352). ¿Podríamos decir que las identidades y diferencias que perciben los catadores profesionales en los aromas y sabores del vino no cuentan como diferencias en sabor o aroma, debido a que no pueden considerarse “normales”? La tesis disposicional no permite dar una definición de *color*, *sabor* o *aroma* independiente de las disposiciones del sujeto percipiente normal. Pero puede mostrarse que hay un análisis de estos términos que trasciende lo disposicional, de modo que (2) no es ni necesario ni suficiente para (1). “A surface that is made up of tiny red and green dots looks yellow to normal observers under normal lighting conditions, but is really red and green as one can verify by examining the surface with a magnifying glass. A circular disk that is half -red and half-green looks yellow when it is spinning to normal observers under normal lighting conditions. But such a disk does not become yellow when it spins” (1982, pág. 353). La tesis disposicional confunde el *parecer inmediato* de sujetos normales en condiciones normales de luz con el color efectivo examinado de manera más detenida en las mismas condiciones. Por otra parte, la tesis de la dependencia parece implicar (al igual que la tesis de la disposicionalidad) que un color (olor, sabor) sólo puede definirse a partir de la naturaleza de *un* sistema perceptual determinado, el humano. En la segunda parte analizaremos las implicaciones de una definición general de color que contenga componentes antropocentristas. Por ahora se señalará que, si la dependencia-de-un-sistema-perceptual es una característica propia de las cualidades

secundarias, entonces no se ha dado un criterio de distinción entre cualidades primarias y secundarias, pues el sistema perceptual, como hecho del mundo externo, ha de tener en sí mismo cualidades primarias. En otras palabras, si una cualidad secundaria es un conjunto de cualidades primarias conocidas imperfectamente, y si se pretende que la adición de “sistema perceptual” explique el específico contenido de una idea producto de cualidades secundarias, entonces es necesario explicar qué tipo de componente agrega el sistema perceptual. ¿Tiene el sistema perceptual (la retina, los conos, los bastones, el nervio óptico, las neuronas) cualidades primarias *que intervienen activamente en la producción de una determinada idea de color?* O, en términos mecanicistas, ¿puede explicarse el funcionamiento del sistema perceptual en sí mismo mediante la postulación de corpúsculos que se mueven por impulso y choque? Si la respuesta es “sí”, entonces la tesis de la dependencia no puede marcar la diferencia entre tipos de cualidades (pues solo tenemos una explicación mecánica de la producción de ideas provenientes cualidades secundarias: esto es, una explicación que se adecua al modo de ser de las cualidades primarias, pero no al modo de ser de las ideas producidas por cualidades secundarias). Si la respuesta es “no”, entonces la interposición de un sistema sensorial entre una mente y un conjunto de cualidades primarias produce un resultado misterioso e inexplicable.

No parece haber dificultad en admitir que las ideas provenientes de cualidades secundarias necesitan de un sistema perceptivo para constituirse como tales, aun cuando el tipo de “manipulación activa” de dicho sistema no quede completamente aclarada. Pero parece necesario indagar cómo es posible afirmar que, en lo que respecta a las ideas provenientes de cualidades primarias, *hay algo más* además de lo percibido. En efecto, dado que el único contacto que tenemos con la realidad es a través de las ideas, resulta aventurado suponer que hay algo más allá de las propias ideas (y esta es justamente la crítica de Berkeley); se necesita de un argumento filosófico para establecer ese *más allá*, puesto que en la idea sólo puede haber idea. Quizás tanto las cualidades secundarias como las primarias puedan explicarse mediante dos tesis:

- d) La tesis corpuscular.
- e) La tesis de la convergencia geométrica.

1.5.1. La tesis corpuscular

Analicemos d). Como hemos dicho, Locke está fuertemente comprometido con el mecanicismo, la idea cartesiana de que todo lo que vemos está compuesto de corpúsculos, y que en rigor son esos corpúsculos los que poseen las cualidades primarias. El movimiento de los corpúsculos provoca todos los sucesos en el mundo, incluso la sensación: tanto el aroma como el color son producto del contacto directo de los corpúsculos con el órgano sensorial. Si a las tres tesis examinadas les incorporamos la tesis corpuscular es posible afirmar que las cualidades primarias se identifican o dependen directamente de los procesos corpusculares de la materia, mientras que las cualidades secundarias dependen sólo indirectamente de ellos, y más directamente de la interposición del sistema perceptual entre el objeto y la mente. Podemos agregar entonces una cláusula a cada una de las tesis expuestas más arriba, para especificar cuál es la relación que tienen estas tesis con la tesis mecanicista corpuscular:

a*) *Tesis de la dependencia*: Mientras las ideas producidas por cualidades secundarias dependen de un sistema perceptivo, las ideas provenientes de cualidades primarias dependen de procesos corpusculares. Dado que las cualidades secundarias son cualidades primarias imperfectamente conocidas, y dado que los sistemas perceptivos también han de poseer cualidades primarias, tanto las cualidades primarias como las secundarias dependen de procesos corpusculares.

b*) *Tesis de la ceguera*: Las ideas provenientes de cualidades primarias poseen en sí mismas las distinciones propias de los procesos mecánicos de la materia, y por lo tanto la ceguera a uno de esos procesos implica la ceguera para cualquier proceso mecánico. La ceguera a una cualidad secundaria sólo está limitada al sistema perceptual, dada la interposición del sistema perceptual entre la realidad exterior y la mente.

c*) *Tesis disposicional*: las ideas provenientes de cualidades secundarias no pueden analizarse exhaustivamente en términos de disposiciones perceptuales, aunque sí pueden analizarse tomando las disposiciones perceptuales *más* el estado de los corpúsculos en el momento de la percepción.

Con estas cláusulas, podemos enlazar la percepción de ideas de cualidades primarias e incluso de secundarias con los procesos mecánicos de los corpúsculos.

Sin embargo, ¿por qué debemos aceptar el mecanicismo corpuscular? La objeción de Berkeley todavía tiene toda su fuerza: que hay corpúsculos de materia es una hipótesis no respaldada por nuestras ideas. Lisa Downing (1998, pág. 381) afirma que la tesis corpuscular es una hipótesis, una ilustración o modelo que puede ayudar a clarificar la noción de *sustancia* aristotélica. “It is clear that he is not here taking it as a starting point but is rather making use of it in order to clarify his conception of real essence” (1998, pág. 385). Sin embargo, no puede ser una *mera* ilustración, dado que el mismo Locke afirma que se trata de la “opinión más racional”: “This question is tied to the question for why Locke characterizes the mechanist conception of real essences as the ‘more rational Opinion’. What is the basis for Locke’s claims about the superiority of the corpuscular hypothesis?” (1998, pág. 385). Las nociones de esencia nominal y esencia real juegan aquí un importante papel: mientras los aristotélicos utilizan la palabra “esencia” para nombrar algo oscuro y confuso³², en Locke la apelación a la teoría corpuscular permite distinguir dos tipos de esencias. La *esencia real*, insensible, desconocida, es la portadora de las cualidades primarias y la causante indirecta de las ideas de cualidades primarias. La idea de esencia en nuestro intelecto es la *esencia nominal*; una construcción hecha a partir de las cualidades que percibimos (primarias y secundarias), construida sobre la base de regularidades perceptuales. Los conceptos de “perro” o de “oro” se refieren a esencias nominales, puesto que las cualidades asignadas a esos conceptos dependen de la manera en que regularmente se nos presentan. Pero no podemos saber si, al margen del modo en que se nos aparecen, esas cualidades son exhaustivas. Tampoco es posible conocer de qué manera se corresponde esa apariencia con las cualidades últimas de los corpúsculos (Un objeto

³² “Locke’s most significant remark in this regard is that the Aristotelians use the Word essence for ‘they know not what’”. (Downing, 1998, pág. 386)

de forma cúbica no está compuesto de corpúsculos cúbicos; diremos más sobre las cualidades y la relación entre lo macroscópico y lo microscópico en 1.7), pero la noción de “corpúsculo” permite anclar la referencia de la esencia nominal en una esencia real, aunque desconocida. “Perro” significa “el conjunto de cualidades que provoca una idea unitaria de cierto tipo”, no significa “la esencia de un conjunto de cuerpos”, porque la esencia de lo corpuscular no se nos puede presentar. Por esto, el *mecanicismo corpuscular* está atado a la noción de *esencia real*: el movimiento de los corpúsculos es parte de la forma de ser de la naturaleza, aunque esta forma sea desconocida y todo lo que conozcamos sea el conjunto de esencias nominales. El mecanicismo corpuscular provee una vía racional para entender las relaciones idea – percepción – cuerpo, y en ese sentido se convierte en *la mejor* hipótesis ilustrativa. Pero, como hipótesis, podría resultar falsa y el escéptico berkeleyano seguiría manteniendo que no existen más que ideas y percepciones, sin materia. Para seguir aceptando la legitimidad de la distinción entre cualidades, es necesario admitir otra tesis.

1.5.2. La tesis de la convergencia geométrica.

Aun si el mecanicismo corpuscular fuera falso, es posible aceptar la distinción entre cualidades a partir de un modelo geométrico. La esencia real manifiesta sus propiedades de modo masivo, continuo y coherente, acorde con un modelo geométrico de la percepción. Una idea proveniente de una cualidad primaria es coherente con otras ideas semejantes. Todo lo que esté presente en nuestro campo sensorial tiene –independientemente de cuál sea su naturaleza última- las propiedades de la forma, la figura, el movimiento, el volumen (o tamaño). La propia idea de superficie incluye la de extensión y tamaño. La idea de una esfera roja en movimiento contiene en sí una coherencia y co – presencia de cualidades primarias tal que, si se le quita una de ellas, no es posible pensar la idea. ¿Qué sería una esfera roja sin la cualidad de la figura?³³ En lo que respecta a la propia percepción, las leyes de

³³ Recordemos lo que decía Bennett: “Las cualidades primarias son aquellas que un cuerpo no puede perder” (Bennett, 1988, pág. 123)

perspectiva remiten a la convergencia de las ideas provenientes de cualidades primarias. No sólo las cosas se mueven; también nosotros podemos movernos y las ideas se nos presentarán como un continuo coherente a través de nuestro trayecto. Si observamos la esfera roja mientras nos movemos, podremos ver cómo se modifica gradualmente la idea inicial, vista a cada paso desde distinta perspectiva, con matices de luz diferentes y en diferente relación con su fondo. Una vez más, no importa cuál sea la naturaleza última de la realidad; la sola *coherencia, coocurrencia y continuidad* (los requisitos C-C-C) de las ideas provenientes de cualidades primarias parecen indicar fiablemente la presencia de algo más allá de nuestras propias ideas. “Cuerpo”, de acuerdo a esta definición, no debe entenderse como una “cosa material”, sino como una función geométrica en la que concurre un cierto tipo de cualidades que satisfacen los requisitos C-C-C. De este modo, las cualidades pueden distinguirse de acuerdo a sus propiedades geométricas o espaciales. Mientras las cualidades primarias son propiedades geométricas, las cualidades secundarias *tomadas en sí mismas* carecen de estas propiedades.

Sin embargo, la tesis de la convergencia geométrica es un argumento débil. Como señalamos en 1.2.3, es posible que se cumplan los requisitos C-C-C de manera puramente subjetiva: todo podía presentarse geoméricamente coherente y, sin embargo, podría no haber absolutamente nada externo que justifique esta presentación. Para que la tesis de la convergencia geométrica alcance sustento, debe complementarse con la tesis corpuscular, o con alguna otra tesis que cumpla un rol epistémico equivalente. Si esta última falla, la primera no puede tener éxito en establecer la diferencia entre ambos tipos de cualidades.

1.5.3. ¿Cómo entender la semejanza (*resemblance*)?

El modelo geométrico de la percepción no puede bastar para caracterizar la diferencia entre cualidades primarias y cualidades secundarias. Por ello, es necesario aceptar alguna tesis que permita anclar los requisitos C-C-C en el mundo externo. El mecanicismo corpuscularista es un candidato, aunque podría no ser el único. De

hecho, la afirmación de que las ideas provenientes de cualidades primarias *semejan* dichas cualidades sólo puede entenderse en el sentido de que hay algo (un objeto externo con ciertas características) a lo que se asemeja. Pero, ¿por qué aceptar que las ideas de cualidades primarias semejan una cualidad en el objeto?

Jonathan Bennett prefiere entender la noción de semejanza de un modo metafórico. “Puesto que las ideas no pueden asemejarse a los cuerpos, esto o bien se descarta o se transforma” (1988, pág. 142). La transformación que propone es: al decir que una idea de figura semeja a la figura, estamos diciendo que usamos el mismo término para nombrar la idea como para nombrar la propia figura; la misma palabra para denominar la causa como el efecto. Esto es plausible, porque hay *identidad* entre las cualidades primarias y el objeto (Bennett, 1988, pág. 136) aunque, desde luego, no hay tal identidad entre la *idea* y el objeto. Pero es posible, también, que Locke pretendiera establecer algún tipo de identidad entre estos últimos. De otro modo, ¿cómo puede una idea contener un *indicio* acerca de su causa? O, dicho de otra manera, ¿cómo podríamos establecer que una idea (el efecto) es *idéntica* a su causa, si sólo contamos con la idea (el efecto) y la causa es sólo inferida? Michael Jacovides (1999, pág. 468) argumenta en favor de una posición literalista. Las ideas tienen una semejanza *literal* con la cualidad, aunque no *estricta*. Si fuera estricta, la semejanza debería ser completa, punto por punto. Pero al ser literal no estricta, hay una semejanza parcial, en algún aspecto, aunque no en todo. En II, VIII, § 15, Locke escribe: “Las ideas de cualidades primarias de los cuerpos son semejanzas (*resemblances*) de ellos, y sus patrones existen realmente en los cuerpos mismos”. Jacovides enfatiza: “A pattern is an exemplar from which copies are supposed to be made. Not every aspect of the pattern is usually supposed to be copied, but at least one attribute should belong to the exemplar and the imitation, at least approximately” (1999, pág. 469)

Si aceptamos la posición literalista de Jacovides, de que en la idea está presente un aspecto del patrón, entonces es posible enfrentar las críticas del escéptico berkeleyano. Se puede aducir que en las ideas hay una huella de correlación entre lo que está presente en la exterioridad y en la propia mente, aunque esa correlación no tiene por qué tener una semejanza total, aunque sí estricta. Si aceptamos la escisión entre realidad mental de la idea y realidad física de los cuerpos, esta semejanza podría

tomar la forma de un isomorfismo parcial, en el sentido estricto de la palabra (no en el sentido metafórico), por ejemplo en relación con la forma geométrica: la idea de un cubo (o el cubo-en-idea) se corresponde con el cubo real (o cubo-en-la-realidad) como se comporta una proyección geométrica en perspectiva de un cubo: hay isomorfismo (parcial) entre el cubo y el cubo proyectado. La mente no puede imitar a la perfección o en su totalidad al objeto pues, en ese caso, ocurriría tal como Sócrates le reprocha a Cratilo con relación a la capacidad de mimesis del lenguaje (Platón, 2009, pág. 52): “todo sería doble”. Esta objeción que Platón realiza a la mimesis del lenguaje, podríamos hacerla a la capacidad representativa de una idea y el hecho causal que la genera: si una idea representa exactamente un hecho externo, entonces no habría manera de distinguir entre ambos.

1.6. La desemejanza de las cualidades secundarias y las ideas: apelación a la arbitrariedad divina.

En 1.5 expusimos dos tesis que sirven de punto de partida para establecer la semejanza entre un cierto tipo de ideas (figura, forma, tamaño, posición) y la realidad externa. Si se acepta la existencia de corpúsculos, o si se entiende como suficiente la tesis de la convergencia geométrica, es posible proyectar la semejanza entre esos corpúsculos / convergencia geométrica y las ideas. Dimos argumentos para mostrar la insuficiencia de ambas tesis: que haya corpúsculos es una hipótesis, y que haya convergencia geométrica de la percepción no garantiza la externidad de lo percibido. La noción de semejanza, a su vez, podía entenderse o bien de manera metafórica, o bien de manera literal. La posición literalista está en mejores condiciones para enfrentar una crítica idealista, pues al afirmar que las ideas provenientes de cualidades primarias tienen ellas mismas, como contenido de conciencia, la cualidad que las originan (la idea de rectángulo tiene en sí misma las cualidades primarias del rectángulo que percibo), se convierte en un candidato más seguro para establecer la conexión entre el aspecto fenoménico y el mundo exterior. Nuestra intención es encontrar los puentes reductivos entre los dos ámbitos, y la semejanza literal entre cualidades primarias e ideas parece prometedora.

Pero aun si se acepta que en la idea proveniente de cualidades primarias está presente, literalmente, un “patrón” de la cualidad instanciada, esto no da ninguna luz con respecto al modo como se dan las ideas provenientes de cualidades secundarias. Si la semejanza es literal no estricta en lo que respecta a las ideas primarias, ¿qué tipo de relación existe con respecto a las secundarias? ¿Se tratará de una semejanza no estricta? ¿O una desemejanza? ¿Existirá una ley que gobierne la aparición de “sensación de rojo” en la conciencia a partir de mecanismos corpusculares? Dado que, según el mecanicismo puesto como candidato, los cuerpos están constituidos por diminutos corpúsculos inobservables, no es posible conocer sus propiedades de manera exhaustiva. La naturaleza última de la realidad está fuera del alcance humano. Por eso, preguntar por la esencia real de las cosas es dar un salto en el vacío: sólo tenemos acceso a las esencias nominales, esto es, ideas abstractas conformadas por conjuntos de ideas cuyas cualidades primarias y secundarias se presentan con regularidad. Para tomar el ejemplo dado más arriba, el término “oro” se refiere no a la esencia real del oro (al *cuerpo* del “oro”), sino al conjunto de cualidades (primarias y secundarias) que en la mente da origen a las ideas asociadas con el oro. La referencia primaria de un término de clase natural es un conjunto de ideas simples, en tanto que esas ideas son producto de un conjunto regular de cualidades.

Como afirma Williams (1969, pág. 319), la noción de *esencia real*, cuyo punto de partida es el mecanicismo corpuscular, requiere de la aceptación de una constante (la constante Φ ³⁴) que no puede ser objeto de pensamiento: no está en lugar de una idea, ni puede jugar un rol psicológico; no puede definirse ostensivamente, ni depende de cómo una cosa se vea, suene, huela o se sienta en un determinado tiempo y espacio: Φ no posee conexión con fenómenos empíricos ni sensoriales de ninguna clase. Un empirista estricto debe rechazar que Φ tenga referencia (como es el caso de David Hume). Sin embargo, en el caso de Locke, la aceptación de la teoría corpuscular como “la opinión más racional” lleva a aceptar que Φ tenga una referencia, aunque esta sea desconocida. Las co – ocurrencias de determinadas cualidades primarias en un

³⁴ Esa constante se reemplaza unívocamente por la noción de *cuerpo*. Lo que para Locke es una constante, en la actualidad sería una variable porque remitiría no al concepto unívoco de cuerpo, sino a clases de cuerpos (electrones, protones, etc.).

determinado espacio y tiempo, se convierten en criterios para inferir la existencia de Φ s y para conocer algunas de sus propiedades. Las cualidades secundarias, en cambio, dado su componente relacional y disposicional, no permiten esta inferencia, al menos de modo directo. Como hemos dicho más arriba, el color de la superficie de un cuerpo –entendido como producto de una cualidad secundaria- es el modo en que la mente resuelve la imposibilidad de formarse una idea de cada forma, figura, posición y movimiento de los corpúsculos que conforman ese cuerpo. Hay, entonces, una conexión entre corpúsculos y cualidades secundarias y en ese sentido la cualidad secundaria tiene un carácter objetivo. La estructura sensorial humana resuelve la imposibilidad de ver las cualidades primarias de los corpúsculos “coloreando” las superficies vistas, haciendo que la superficie se vea roja o verde. Podría pensarse que esta estructura conforma el costado *subjetivo* de la percepción. Sin embargo no es así: para Locke, que veamos rojo en una determinada superficie no es un capricho azaroso; no existe la posibilidad de que otra persona en circunstancias similares (y cuya estructura sensorial sea normal) vea un color diferente. En II, XXXII, § 15, Locke ensaya la posibilidad de que dos hombres tuvieran una indistinguible diferencia en la estructura de sus órganos sensoriales, lo que podría provocarles ideas diferentes acerca de un mismo hecho:

(...) Por ejemplo, que la idea que produjera una violeta en la mente de un hombre por conducto de su vista fuese la misma idea producida en la mente de otro hombre por una caléndula y viceversa. Porque, como esto no podría jamás saberse, ya que la mente de un hombre no podría pasar al cuerpo del otro, a fin de percibir qué apariencias se producirían por esos órganos, ni las ideas así formadas, ni los nombres que las denotan tendrían confusión alguna, ni habría falsedad en las unas y en los otros; porque, como todas las cosas que tuvieran la misma textura de una violeta producirían de un modo constante la idea que uno de esos hombres denominaría azul, y aquellas que tuvieran la textura de una caléndula producirían de un modo constante la idea que constantemente ha denominado amarillo, fueren cuales fueren las apariencias que tuviere en su mente, podría distinguir con igual constancia, por su uso, las cosas que tuvieran esas apariencias, y podría también entender y dar a entender esas distinciones señaladas por las palabras azul y amarillo”, como si las ideas en su mente, recibidas de esas dos flores, fueran exactamente las mismas que las ideas recibidas por las mentes de otros hombres. (Locke, [1710] 1956, pág. 375)

Supongamos que los dos hombres de este ejemplo tienen una estructura sensorial funcionalmente normal y que estén bajo idénticas condiciones de luz. Planteado así,

estamos ante el problema de los *qualia* invertidos. ¿Qué nos dice Locke acerca de la posibilidad de este fenómeno? En primer lugar, que A vea amarillo donde B ve azul (y viceversa) no provoca ningún problema semántico, ni malentendidos de ninguna clase. ¿Cómo es posible esto, si la referencia de A es una idea de amarillo, y la de B es una idea de azul? El punto es que los términos “azul” y “amarillo” parecen anclar su referencia no en las ideas de azul y de amarillo en nuestra mente, sino en la cualidad secundaria que es idéntica en ambos casos: los poderes causales de la flor que provocan la idea de amarillo en A son idénticos a los poderes causales que provocan la idea de azul en B. La idea de amarillo y la idea de azul son efectos de esos poderes causales. Como los efectos (las ideas) son privados, no hay posibilidad de error, y se puede utilizar sin problemas el lenguaje, dado que no se lo utiliza para referirse a las propias ideas, sino a las cualidades. B nunca podrá saber si esa cualidad secundaria que invoca como causante de su idea de “azul” está provocando en su mente la idea que para A es, en realidad, de “amarillo” y viceversa, pero en rigor tal inversión es imposible, porque no hay espacio de diferencia subjetiva para que una cualidad provoque una idea diferente en cada sujeto.

Existe, sin embargo, un problema con esta afirmación de imposibilidad: el único criterio de que disponemos para reconocer una cualidad secundaria es la idea que se forma en nuestra mente. Si dos personas se forman ideas cualitativamente distintas, no se garantiza la identidad de la cualidad. Quizás la cualidad A* que produce la idea de amarillo en A sea diferente de la cualidad B* que provoca azul en B. Quizás A sea ciego a la cualidad B*, y B sea ciego a la cualidad A*. ¿Quién puede garantizar que el poder causal de producir la idea de amarillo no está presente en una violeta? Si las disposiciones sensoriales de toda la humanidad cambiaran e hicieran que las violetas nos provocaran ideas de color amarillo, diríamos que el amarillo es producto de un poder causal de las violetas. El poder causal parece *independiente* de la idea que nos formemos, pues tal idea está condicionada por la estructura perceptiva.

Sin embargo, Locke desecha la posibilidad de que dos hombres pudieran tener ideas cualitativamente diferentes a partir de una misma cualidad.

(...) Yo me inclino mucho a pensar que las ideas sensibles producidas por cualquier objeto en la mente de diferentes hombres son, por lo común, muy cercana e indiscerniblemente parecidas. (Locke, [1710] 1956, pág. 375)

No ofreceré un argumento ulterior, pero la posibilidad de los qualia invertidos se cierra con estas palabras y no vuelve a tratarse en el resto del *Ensayo*. Hacking (1979, pág. 65) comenta al respecto: “Se sigue de todo esto que los problemas acerca de las ideas de violetas y caléndulas tienen muy poca importancia para la teoría del conocimiento del siglo diecisiete”. La doctrina mecanicista – corpuscular implica que la sensación se produce a partir de partículas que golpean las terminaciones nerviosas. “Locke es muy poco claro al respecto de cómo esos impulsos físicos producen las ideas de color, aunque sin duda su producción responde a algo como leyes. A menos que supongamos fundadamente que la fisiología y la psicología de otra persona difieren de la mía, podemos pensar que es probable que las mismas leyes de la física y de la psiquis valgan en ambos casos y produzcan efectos similares a partir de causas similares” (1979, pág. 65).

Aun cuando aceptemos la plausibilidad que Hacking supone, Locke explícitamente excluye la posibilidad de los qualia invertidos, para la cual no es necesario postular una diferencia nomológica en la fisiología y la psicología. Que causas similares produzcan efectos similares no parece discutible en el terreno de las cualidades primarias. Pero – agregaríamos nosotros- se necesita de un argumento filosófico para aceptar esto acerca de las cualidades secundarias. Porque las ideas que son producto de las cualidades secundarias no sólo surgen de la estructura sensorial puramente corporal: también son creadas y modeladas espontáneamente por la mente, en el sentido de que no hay semejanza entre la causa, la estructura sensorial y la idea. Mientras el poder causal del cuerpo rojo y la estructura sensorial pueden explicarse mediante pautas puramente mecánicas, la *modalidad* de la idea producida es sólo producto espontáneo de la mente. ¿Por qué vemos a las cosas rojas y olemos a las cosas fragantes, y no al revés? No hay contradicción en que los mismos poderes causales nos provean una idea visual de los aromas, y una idea olfativa de los colores. ¿Qué hace que esas modalidades tengan la cualidad específica que poseen, y no puedan intercambiarse? Locke no puede aducir más que correlatos entre la estructura

perceptiva y la idea en la mente, pero el modelo mecanicista no puede explicar por qué las cosas nos aparecen del modo en que nos aparecen. El mismo Locke asume las limitaciones del modelo mecanicista. En IV, III, § 29 afirma que hay ciertas ideas que están tan fuertemente relacionadas entre sí que resulta imposible pensarlas separadas:

Así, por ejemplo, la idea de un triángulo rectilíneo lleva consigo necesariamente la igualdad de sus ángulos a dos rectos. Y no podemos concebir que esta relación, esta conexión entre esas dos ideas pueda ser mudable o que dependa de **un poder arbitrario** que la haya establecido a su gusto, o que pueda establecerse de otro modo. Pero en lo que se refiere a la coherencia y continuidad de las partes de la materia; a **la producción en nosotros de las sensaciones de color o de sonidos**, etc., en virtud de un impulso o movimiento; es más como las reglas originales y la comunicación del movimiento son tales que no podemos descubrir en ellas ninguna conexión natural con ninguna de las ideas que tenemos, **no podemos menos de adscribirlas a la voluntad arbitraria y al beneplácito del Sabio Arquitecto del universo** (Locke, [1710] 1956, pág. 558, cursivas añadidas)

La cita anterior parece entrar en conflicto con la conclusión de Hacking expuesta más arriba: la conexión entre las cualidades secundarias y las ideas producidas parece producto de una arbitrariedad divina no sujeta a leyes mecánicas ni de ningún tipo. De esto último podemos establecer una curiosa conexión semántica entre los términos que se refieren a ideas producidas por cualidades secundarias. Puede decirse que el término “rojo”, cuando se lo usa en sentido particular (para nombrar a “el rojo que estoy viendo en este momento”) se refiere, en sentido primario, a mi idea de rojo. Si queremos referirnos no a la idea particular de rojo, sino a la esencia nominal, la referencia es el conjunto de cualidades secundarias que conforman la idea abstracta de “la rojez”. Pero en una tercera acepción puedo referirme al modo en que la cualidad secundaria se conecta con mi idea, y en ese caso “rojo” será “lo que el Arquitecto Universal haya decidido para unir la cualidad secundaria con mi idea”. En esta acepción, “rojo” equivale a una “x”, una incógnita no sujeta a leyes mecánicas (ni mecánicas, ni de ningún tipo) cuya realización efectiva no necesariamente debe ser igual en dos individuos, por más que las leyes mecánicas que gobiernan su fisiología y su psiquis sean idénticas entre sí. Una vez que se acepta la posibilidad de una anómala intervención divina, todo está permitido. Dios podría hacer que dos personas cuyas

estructuras perceptivas son normales, tengan ideas diferentes a partir del mismo tipo de cualidades secundarias.

La apelación a la arbitrariedad divina es la aceptación de una brecha inconmensurable entre lo mental y lo físico. Es la claudicación de la *legalidad, unidad y reductividad* de los dominios fenoménicos sobre los físicos y resulta, a la postre, contraproducente para todo el sistema mecanicista. Aun si se acepta que las cualidades primarias dejan una huella en la idea mediante cierta relación de semejanza, esto no es ni remotamente posible con las cualidades secundarias, pues la apelación a la arbitrariedad divina es la asunción de que entre los factores causales y la idea hay un elemento no explicable en términos legales mecanicistas y, en rigor, sujeto a la voluntad inescrutable de un ser omnisciente³⁵.

1.7. ¿Toda cualidad es secundaria?

En el apartado anterior concluimos que la apelación a la arbitrariedad divina parece socavar profundamente cualquier intento por encontrar un puente legal entre las cualidades secundarias y las ideas derivadas a partir de ellas. Pero, ¿está tan claro que las ideas provocadas por cualidades primarias tienen algún tipo de semejanza con un hecho externo? La ciencia del siglo XVII asume que, una vez obtenido el registro de todas las cualidades primarias del universo, obtendremos todo el conocimiento posible acerca del mundo, y de allí podremos efectuar una reducción sin resto de nuestras ideas de cualidades primarias a sus causas externas y, por añadidura, sin la adición de ningún principio adicional, habremos explicado también en qué consisten las cualidades secundarias. No podremos explicar – como hemos visto en el apartado anterior- cómo esas cualidades secundarias generan ideas (pues allí puede estar

³⁵ Conviene remarcar que esta tesis de la arbitrariedad con respecto a las ideas derivadas de cualidades secundarias ya se encuentra en Descartes: “Ninguna afinidad ni relación existe (comprensible para mi, al menos) entre la emoción del estómago y el deseo de comer, ni tampoco entre el sentimiento de dolor y los tristes pensamientos que provoca” (*Meditaciones Metafísicas*, sexta meditación, § 5, pág. 57; véase también § 14 y § 21). Leibniz rechazará la idea de que Dios pueda hacer algún tipo de conexión arbitraria, cf.: “Tampoco hay que imaginarse que ideas como el color y el dolor sean completamente arbitrarias y no tengan ningún tipo de conexión o relación natural con sus causas: Dios no tiene la costumbre de actuar con tan poco orden y razón” (Leibniz, Nuevos ensayos sobre el entendimiento humano, libro II, cap. VIII, pág. 145)

presente la arbitrariedad divina, o algún tipo de legalidad divina desconocida, tal como afirmaría Leibniz), pero hay un intento de englobar en un corpus explicativo unitario de qué modo una cualidad secundaria se deriva de un conjunto de cualidades primarias. En este apartado veremos qué tan viable resulta este proyecto desde el punto de vista conceptual.

Williams (1969), asevera que

The seventeenth-century ideal of a scientific account of the universe is just this –a complete record of the position, size and figure of all the elementary particles at every moment of time, a record which could be deduced from a knowledge of the position, size, velocity and figure of these particles at any one moment in time, together with the fundamental principles of Mechanics. (1969, pág. 315)

La hipótesis mecanicista es la siguiente: si se obtiene el registro completo de las cualidades primarias de los corpúsculos, tendremos la totalidad de los elementos a los que deben reducirse las cualidades primarias de los cuerpos macroscópicos y las cualidades secundarias. Todo lo que necesitamos son principios puente para conectar las cualidades primarias macroscópicas y las cualidades secundarias, con las cualidades primarias de los corpúsculos. Parte del ideal científico del siglo XVII está dado por el intento de explicación de las cualidades secundarias a partir de las cualidades primarias. Tanto las cualidades primarias de los cuerpos macroscópicos como sus cualidades secundarias deberían poder explicarse *sin resto* a partir de las cualidades primarias de los corpúsculos, de manera tal que una vez dadas *todas* las cualidades primarias de los corpúsculos, se obtenga la identidad entre las cualidades primarias macroscópicas y las cualidades secundarias con las cualidades primarias de los corpúsculos. Esta es la tesis reductivista estándar. Kim (1998) la caracteriza como

... [A] relation between two *scientific theories*, the target theory that is up for reduction and the theory that serves as its reduction base (“the reducer” or “base theory”). A theory in this context is taken to be a set of statements, which represent its “laws”. (1998, pág. 212)

Para poder establecer esta relación, se necesita de principios-puente que conecten las expresiones de ambas teorías. Los principios-puente deben servir para mostrar la identidad de las propiedades de la teoría a reducir con las propiedades de la teoría base. En el caso que nos ocupa, la reducción podría ser exitosa si se mostrara que las

propiedades de las cualidades secundarias y de las cualidades primarias son idénticas a las propiedades de las cualidades primarias de los corpúsculos, y para ello debemos recurrir a principios-puente que conecten ambos tipos de ámbitos. Por lo tanto, el problema consiste en la posibilidad de obtener esos principios puente.

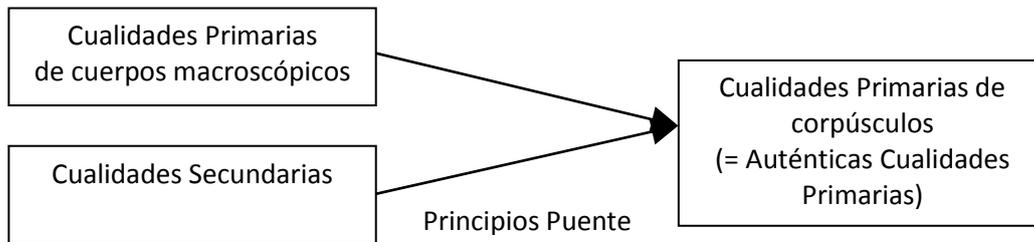


Figura 5

Si la hipótesis corpuscular es *cierta*, entonces es claramente plausible que encontremos principios puente que conecten cualidades macroscópicas con cualidades microscópicas. Pero, como notamos en 1.5.1, la noción de que existen corpúsculos que se comportan de acuerdo a leyes mecánicas es una hipótesis filosófica para tender el puente entre las esencias nominales y las esencias reales, pero no una consecuencia empírica del avance científico del siglo XVII. Los corpúsculos son postulados como “la hipótesis más racional” posible (v. *supra*, 1.5). Pero Locke no se muestra esperanzado en que estas entidades microscópicas puedan conocerse a partir del avance científico, lo que indica un escepticismo con respecto al conocimiento de estos principios puente. Yost afirma al respecto:

Although Locke was aware of the advances that had been made in microscopy, he apparently did not think that one could ever see, more or less directly, the specific corpuscular mechanisms corresponding to observable species. (1951, pág. 124)

Yost considera que diversos pasajes del Ensayo (cfr: §§ II.XXIII.32; II.XXXI.6; III.VI.9; III.X.19; IV.VI.5; IV.XII.11) no parecen afirmar el obvio hecho de que los corpúsculos no son visibles al ojo desnudo, sino algo más: “I think they were intended to assert that corpuscular mechanisms would never be ‘seen’ with an instrument” (Yost, 1951, pág.

124) En III.VI.9 Locke afirma abiertamente la inconmensurable complejidad de las esencias reales y la profunda limitación de nuestros mecanismos cognitivos:

Nuestras facultades no nos conducen, en el conocimiento y distinción de las substancias, más allá de una colección de aquellas ideas sensibles que advertimos en ellas; las cuales colecciones, por más que se formen con la mayor diligencia y exactitud de que seamos capaces, *están más alejadas de la verdadera constitución interna de donde fluyen esas cualidades* que, como dije, lo está la idea de un campesino respecto al mecanismo interior de aquel famoso reloj de Estrasburgo, del cual únicamente advierte la forma y el movimiento exteriores. (...) Cuando nos ponemos a pisar las piedras que pisamos o el hierro que a diario manejamos, pronto advertimos que su factura nos es desconocida y que *somos incapaces de dar razón de las diferentes cualidades que encontramos en esas cosas*; es evidente que su constitución interna, de donde dependen sus propiedades, escapa a nuestro conocimiento. (...) Y, sin embargo, todos saben bien cuán cortos se quedan esos ejemplos, respecto a los finísimos mecanismos e inconcebibles esencias reales de las plantas y de los animales. El artificio empleado por el Omnisapiente y Todopoderoso Dios en la grandiosa obra del Universo y en cada una de sus partes *excede más los alcances y la comprensión del hombre más inquisitivo e inteligente*, que la mejor obra del hombre más ingenioso excede la comprensión de la más ignorante de las criaturas racionales. Resulta, pues, *que en vano pretendemos ordenar las cosas en clases y disponerlas en especies, bajo nombres, en razón de su esencias reales, que tan lejanas están de ser descubiertas y tan remotas están al alcance de nuestra comprensión* (Locke, [1710] 1956, pág. 436, cursivas añadidas)

Con este pasaje, es difícil afirmar que Locke creía que los instrumentos científicos revelarían alguna vez las propiedades de la esencia real. Parece claro que las cualidades primarias que podemos observar *no son las cualidades últimas de la realidad* y que, por lo tanto, las cualidades primarias macroscópicas deberían poder reducirse a las microscópicas. Desde este punto de vista, las Q1 percibidas inmediatamente por los sentidos están tan necesitadas de reducción como las Q2. Si una idea es *literalmente* semejante a una Q1-macro no tiene por qué ser literalmente semejante a una Q1-micro. Los corpúsculos de una figura cúbica o esférica no tienen que ser, ellos mismos, cúbicos o esféricos. En el caso de una idea de una figura esférica, la semejanza entre la idea y la cualidad primaria sólo opera al nivel macroscópico, pero no en el nivel microscópico. De aquí se infiere que las ideas Q1-macro no pueden ser semejantes a los hechos últimos, o en todo caso sólo pueden serlo de un modo sinuoso e indirecto, y en esto las Q1 macro son tan pasibles de una

reducción como las propias Q2. Presumiblemente, hay cualidades primarias corpusculares irreductibles, pero no hay prospecto de que sean cognoscibles y mucho menos, entonces, hay prospecto de conseguir los principios puente que conectarían ambos terrenos. Los tipos de reducción operadas sobre las cualidades serían los siguientes:

- a) Toda Q2 de cualquier nivel debe poder reducirse a una Q1 de un nivel inferior.
- b) Es debatible si en el nivel mínimo de la materia pueden encontrarse Q2 (Puede ser un error categorial pretender encontrar cualidades secundarias en el último escalón de la materia, o bien, como afirmaba Berkeley, es plausible que se encuentren también en ellas)
- c) Las Q1 de un determinado nivel deben poder reducirse a Q1 de un nivel inferior, aunque:
- d) No toda Q1 puede reducirse a una Q1 de nivel inferior: hay Q1 micro que representan el nivel mínimo de la materia.
- e) No contamos de hecho con principios-puente en todos los casos en que es necesaria la reducción, de modo que dicha reducción sólo es un desiderátum.

Si leemos en el Ensayo, § II.VIII.17: “El volumen, el número, la forma y el movimiento particulares de las partes del fuego o de la nieve están realmente en esos cuerpos, sean o no percibidos por los sentidos de alguien, y por eso puede llamárseles cualidades reales, porque realmente existen en esos cuerpos” (pág. 116), junto con § II.VIII.24 donde se afirma que las cualidades primarias son semejanzas, podemos notar que las cualidades *reales* son *partes* (presumiblemente insensibles) de los cuerpos. Esto parece indicar claramente que las cualidades primarias que se perciben inmediatamente (esto es, las macroscópicas) no cuentan como cualidades *reales*, sino que dependen de otras cualidades subyacentes. De aquí podemos inferir:

- f) Las cualidades primarias y secundarias de cualquier nivel pueden reducirse a las cualidades *reales*, esto es, cualidades primarias de los corpúsculos.
- g) Estas cualidades reales son desconocidas y -en apariencia- inasequibles al entendimiento humano.

Dada la imposibilidad de conocimiento de las cualidades *reales*, Locke recomienda el *método histórico*. Según Yost:

Whenever he spoke of the methods of increasing empirical knowledge, he recommended the “historical” method, not the “speculative” or “hypothetical” method (1951, pág. 127)

Este método histórico busca principios – puente acerca de correlaciones entre cualidades observables, pero no acerca de lo inobservable. En § IV.XII.10 Locke argumenta:

No niego que un hombre habituado a realizar experimentos racionales y regulares no sea capaz de penetrar más en la naturaleza de los cuerpos, y conjeturar con mayor tino acerca de sus propiedades desconocidas, que uno que no tenga esa familiaridad; pero, con todo, esto no es, según ya dije, sino juicio y una opinión, y no conocimiento y certeza. Esta manera de adquirir y de adelantar nuestro conocimiento acerca de las substancias, es decir, sólo por medio de la experiencia y de la historia, que *es cuanto permite la flaqueza de nuestras facultades en este estado de mediocridad en que nos encontramos en este mundo, me hace sospechar que la filosofía natural no es capaz de constituirse como ciencia*. Pienso que podemos aspirar a tener un conocimiento *general* muy escaso acerca de las especies de los cuerpos y sus diversas propiedades. Podemos, en efecto, tener experiencias y observaciones históricas de donde es posible sacar ventajas de bienestar y salud, e incrementar de esa manera nuestro acervo de comodidades para la vida; pero más allá de esto, *mucho me temo que nuestro talento no pueda pasar, y que nuestras facultades, según conjeturo, no puedan avanzar*. (Locke, [1710] 1956, págs. 648-9, cursivas añadidas)

Este pesimismo teórico está combinado con un moderado optimismo pragmático: la ciencia nos conduce a un control de la naturaleza y a una mejora en las condiciones de vida, siempre y cuando el científico o el naturalista no intenten penetrar la esencia real y no se embarque en la búsqueda de los principios- puente que conectan lo observable con lo inobservable. En § IV.XII.12 dice “Respecto al conocimiento de los cuerpos, debemos conformarnos con adivinar lo que se pueda a base de los experimentos particulares, ya que no podemos, partiendo del descubrimiento de sus esencias reales, aprehender de un golpe colecciones enteras, y en conjunto comprender la naturaleza y las propiedades de toda una especie.” ([1710] 1956, pág. 650).

Es importante tener en cuenta que este rechazo de las hipótesis sobre inobservables es una consecuencia del programa empirista llevado a cabo por Locke, y por (quizás de

un modo práctico más fundamental) por sus contemporáneos Boyle y Sydenham ³⁶. En el programa empirista de Locke, los argumentos hipotéticos deben estar vinculados a la experiencia, y la única regla de la experiencia admisible para obtener conclusiones meramente probables es la analogía. Las hipótesis sobre inobservables no pueden basarse en principios innatos, pues de haber tales principios no pueden estar impresos naturalmente en el alma, porque son desconocidos por los niños y por los idiotas (Ensayo, § I, II, 5), de modo que no puede existir un principio a priori al que podamos recurrir para investigar aquellas cosas que están fuera del alcance de los sentidos. La analogía es un procedimiento legítimo y compatible con el tratamiento empírico de la naturaleza. La analogía fundamental es la que presupone la continuidad entre los hechos macroscópicos y los hechos microscópicos:

En las cosas que no pueden ser descubiertas por los sentidos, la analogía es la gran regla de la probabilidad. Las probabilidades que hemos mencionado hasta aquí no se refieren sino a asuntos de hecho y de cosas capaces de ser observadas y testimoniadas. [...] Así, observando que la pura fricción entre dos cuerpos produce calor, y muchas veces llega a producir hasta el fuego mismo, tenemos motivos para pensar que lo que llamamos calor y fuego consiste en una agitación violenta de las partículas imperceptibles de la materia incandescente (§ IV, XVI, 12, pág. 670)

La analogía, entonces, se basa en la continuidad de la experiencia y en la noción de ascenso y descenso graduales de todo lo que existe:

Así, encontrando que en todas las partes de la creación, que caen bajo la observación humana, existe una conexión gradual entre unas y otras, sin ningún vacío considerable o discernible entre ellas, entre esa gran variedad de las cosas que vemos en el mundo, que están tan estrechamente unidas entre sí, que en los diversos rangos de seres no es fácil descubrir los límites entre ellos, tenemos motivo para quedar persuadidos de que las cosas, paso a paso, ascienden hacia la perfección por grados insensibles (§ IV, XVI, 12, pág. 670-1)

³⁶ En lo que atañe al argumento hipotético por analogía fundada en la experiencia, como en el tratamiento de las nociones de cualidad primaria y secundaria, Locke no difiere de Robert Boyle (1627 – 1691). “La distinción en ambos autores se basa en unos presupuestos tan similares y resulta tan parecida conceptual y terminológicamente, que no resulta nada descabellado que Locke la tomara *tal cual* de las obras de Boyle” (Lorenzo, 1996, pág. 112). Thomas Sydenham (1624 – 1689), médico, colega de Locke, preconizó un método de observación de las enfermedades sin la intervención de especulaciones y rechazaba toda explicación de las enfermedades que apelara a cualidades, fuerzas o mecanismos ocultos. Todo lo que tenía que hacer la medicina era clasificar semiológicamente las enfermedades, por sus propiedades observables. John Locke admiraba a ambos y confiesa haber trabajado, literalmente, como “simple obrero” de ellos (y de Newton y Huygens) “en la tarea de desbrozar un poco el terreno y limpiarlo de los escombros que entorpecen la marcha del saber” (Locke, [1710] 1956, pág. 10)

¿Deben rechazarse las hipótesis? En § IV.XII 13 propone que

Las hipótesis, con tal de que estén bien formuladas, son por lo menos grandes auxilios para la memoria, y frecuentemente nos dirigen hacia nuevos descubrimientos. Lo que he querido decir es que no debemos abrazar ninguna con demasiada precipitación (...) hasta no haber examinado con esmero las particularidades del caso, y haber realizado varios experimentos en aquella cosa que pretendamos explicar por medio de la hipótesis, para así ver si contesta debidamente, si nuestros principios nos acompañan en todo, y que no resulten tan incompatibles respecto a nuestro fenómeno examinado ([1710] 1956, pág. 651)

Sin embargo, las hipótesis deben ser *históricas*, esto es, *descriptivas de lo que es dado a los sentidos* y no acerca de hechos microscópicos. No es posible descubrir la *conexión necesaria* entre los eventos observables, sino *correlaciones* entre hechos observables, lo que nos permitía inferir, en general y sin certeza (mediante una analogía fundada en la continuidad de los hechos microscópicos y macroscópicos, cfr. Locke [1710] 1956, IV, XVI, § 12), alguno de los mecanismos subyacentes del proceso observado. Esta afirmación nos deja sin ningún tipo de principio puente: ni de lo macroscópico a lo microscópico, ni de lo microscópico a lo macroscópico. En una teoría reductivista los principios-puente nos deberían mostrar que la propiedad A es, en realidad, idéntica a una propiedad basal A*. El método histórico, en cambio, no está en condiciones de afirmar identidades sino correlaciones.

Aun cuando se afirme que las cualidades secundarias se reducen a las primarias, y que las cualidades macro se reducen a las cualidades micro, el método de la ciencia no debe consistir en hipótesis acerca de los principios-puente que conectan estos ámbitos. Que el oro se disuelva en agua regia es un hecho histórico que forma parte de la descripción de la esencia nominal del oro. Pero cuál sea la conexión entre este hecho histórico y la esencia real, es algo que no debe siquiera indagarse. De este modo, **el reduccionismo de Locke no forma parte de un programa científico, sino de un programa filosófico: que las cualidades sean reducibles a conjuntos de cualidades reales** (cualidades últimas de los corpúsculos) **es una tesis metafísica y no un principio guía para la investigación científica**. Las cualidades reales y las esencias reales están allí, afuera. Pero el entendimiento humano no debe perder el tiempo buscándolas. De este modo, las cualidades *reales* no juegan ningún rol en el conocimiento humano ni

en la práctica científica: son producto de una hipótesis filosófica acerca de la estructura basal de la realidad, pero tal hipótesis resulta epistémicamente inerte, porque los principios metodológicos recomendados para la práctica científica no pueden utilizar esa hipótesis para conectar los fenómenos observables con los principios mecánicos últimos. Esta situación epistémica es sumamente precaria: es un realismo *de facto* (pues afirma la existencia de los corpúsculos dado que es la “hipótesis más racional posible”), pero un pragmatismo pesimista *de iure*, pues la afirmación de la existencia de corpúsculos es una hipótesis desligada de cualquier recomendación epistémica: la ciencia no debe buscar en los corpúsculos el *explanans* de los fenómenos observables; por el contrario, debe abstenerse de una tarea inquisitiva que pretenda conectar los fenómenos con las cualidades *reales*. El programa reductivista de Locke, entonces, no puede efectivizarse. El mundo está ahí y tal vez sea intrínsecamente reducible, pero la mente humana está incapacitada para ello.

Pero también hay una brecha en lo que respecta a las cualidades y las ideas. Las ideas provenientes de cualidades primarias semejan literalmente a las cualidades que están en los cuerpos. Pero las cualidades macroscópicas no tienen por qué ser semejantes a las microscópicas, y en rigor cada nivel microscópico no tiene por qué ser semejante a su nivel superior. De modo que las *ideas* de cualidades primarias no tienen por qué semejar cualidades *reales*: en todo caso, es remota la posibilidad de que nuestra mente alguna vez pueda tener una idea de una cualidad *real*. Esta imposibilidad termina diluyendo la difusa distinción entre cualidades primarias y secundarias, pues si nuestra mente no percibe cualidades primarias “reales”, sólo puede percibir forma, figura, posición, número y movimiento de eventos en su estructura macroscópica, i.e., tal como se los presentan nuestras disposiciones sensoriales (tesis de la disposicionalidad) y dependen de una estructura sensorial (tesis de la dependencia) (cfr. *supra*, 1.5). Y los argumentos de la disposicionalidad y de la dependencia (junto con el de la ceguera) eran lo que distinguía a las cualidades primarias de las secundarias. En rigor, nuestra percepción solo lo es de eventos macroscópicos, y ninguna de las cualidades dadas en tales eventos son “reales”: los eventos microscópicos subyacentes suceden sin nuestro conocimiento, mudos e inaccesibles.

Las verdaderas cualidades primarias –las cualidades *reales*- están fuera de nuestro alcance. En definitiva, nuestras ideas de cualidades primarias macroscópicas contienen un elemento fenoménico que no es explicable por medio de cualidades primarias *reales*, puesto que no tenemos acceso a ellas: las ideas provenientes de cualidades primarias macroscópicas están tan necesitadas de explicación como las ideas de cualidades secundarias. De modo que a la lista de elementos dados a la reducción deberíamos agregar:

- h) Las cualidades primarias que percibimos de los cuerpos son, en verdad, cualidades secundarias de cualidades primarias insensibles “reales”

Debemos concluir, entonces, que el trayecto explicativo- reductivo entre la idea y sus causas no puede completarse: las ideas provenientes de cualidades secundarias no pueden en rigor reducirse a sus causas; las ideas de cualidades primarias que se nos presentan ante nuestros sentidos no semejan a las cualidades primarias “reales”, microscópicas, subyacentes en los corpúsculos que forman la realidad, y que, por lo tanto, todas nuestras ideas fenoménicas, sean de cualidades secundarias como de cualidades primarias, necesitan anclarse en cualidades primarias “reales” para que la reducción tenga siquiera una mínima posibilidad de efectuarse. Sin embargo, el conocimiento de estas verdaderas cualidades primarias no es asequible. De modo que las nociones de cualidad primaria – secundaria, que debían servirnos para tender el puente entre nuestras ideas y el mundo, se convierten ellas mismas en conceptualizaciones ociosas y en rigor no pueden servirnos como auxiliares conceptuales para lograr una visión unificada del mundo y nuestra conciencia inserta en él. En definitiva, si Locke fuera consecuente con sus propias conceptualizaciones, los puentes tendidos entre las ideas y las esencias reales son tan pequeños y tan débiles que no nos permiten avanzar ni un tramo sobre ellos.

1.8. Frege y su lucha contra el representacionalismo

Hemos examinado que en Locke, como representante del realismo por representación existe una seria dificultad para atribuir un rol decisivo a las cualidades intermediarias

entre las ideas y el mundo externo y, en definitiva, para garantizar un nexo entre “lo interno” y “lo externo”. Debido a que la mediación está anclada en criterios que pone la propia subjetividad, resulta plausible pensar que aun las ideas que provienen de cualidades primarias no tengan semejanza con ningún objeto externo. La posibilidad de esta desemejanza probablemente contribuya a desbaratar el esquema representacionista. A menos que se admitan abierta y acriticamente la veracidad divina y el criterio de certeza con respecto a las cualidades primarias, y la participación arbitraria de Dios para el establecimiento de las ideas derivadas de cualidades secundarias, el realismo por representación parece condenado al fracaso -y la apelación a la divinidad como posible argumento explicativo es la expresión de dicho fracaso. En este apartado examinaremos la postura de Frege en tanto encarna una propuesta correctiva con respecto al realismo por representación. Parece que, si no aceptamos el realismo por representación para explicar la causalidad de las ideas, quedan las siguientes tres opciones: a) un idealismo extremo, con el consiguiente peligro de solipsismo y subjetivismo particular; b) un fenomenalismo, en el cual todo conocimiento está anclado en el fenómeno representable (y la causalidad queda fuertemente cuestionada), o c) un realismo débil que no recurra a la noción de representación para garantizar la aprehensión de una causalidad³⁷. Esta última es la posición de Frege.

³⁷ Omitimos enumerar aquí la posición de la fenomenología clásica, para la cual no tiene sentido el planteo acerca de la causalidad de las ideas. Para la fenomenología, la demarcación entre lo interno y lo externo resulta un modo inadecuado de plantear el problema de la naturaleza de lo mental. Cf., “Todo existente, sea cual fuere el dominio del ser al que pertenezca, es un índice para un sistema de vivencias en las que está “dado” según leyes a priori [...] Husserl alcanza desde el periodo de las *Investigaciones Lógicas* el punto central en que el sujeto y el objeto aparecen como inseparables. La hipótesis de un *en sí* deviene tan absurda como la de una conciencia que no percibiese el mundo *en sí mismo*”. (Tran-Duc-Thao, *Fenomenología y Materialismo Dialéctico*, 1971, Buenos Aires, Nueva Visión, pág. 45) Esta percepción del “mundo en sí mismo” es lo que Merleau – Ponty expone como algo dado a través del *sentir* en un campo fenoménico: percibir, ser consciente-de-algo, es sentir ese algo no como un conjunto de cualidades “como dadas afuera”, muertas, sino activas, con un sentido para el horizonte de la conciencia: “La visión está ya habitada por un sentido que le da una función en el espectáculo del mundo, lo mismo que nuestra existencia. El quale puro solamente nos sería dado si el mundo fuese un espectáculo y el propio cuerpo un mecanismo del que tomaría conocimiento una mente imparcial” (Merleau-Ponty, *Fenomenología de la Percepción*, 1994, Buenos Aires, Planeta - Agostini, pág. 78)

Antes de analizar la posición de Frege, mencionemos brevemente cuáles son los peligros del punto b), el fenomenalismo. La posición fenomenalista no puede efectuar una conexión entre “lo interno” y “lo externo” o, en todo caso, lo “externo” es una construcción de los sujetos. Dicha construcción puede tener validez intersubjetiva, pero está anclada en entidades puramente internas. En palabras de García Carpintero (1996, pág. 155):

El fenomenalista (...) ha redefinido “mundo objetivo” de modo que la relación de “eso” con nuestro pensamiento ya no es *causal* (en el sentido realista de ‘causal’); en palabras de Russell y Wittgenstein, la relación es puramente definicional: las “cosas” son meros “constructos lógicos” a partir de ideas, en lugar de ser entidades que causan las ideas. El “mundo objetivo” es una construcción lógica elaborada haciendo exclusivamente referencia a entidades internas.

Es importante destacar, entonces, que el problema del realismo por representación *persiste* en el fenomenalismo: tanto en uno como en otro caso, no se pueden tender los puentes entre los estados mentales y la realidad y, en definitiva, las “entidades internas” que menciona García Carpintero son decisivas para la construcción de lo que se considere “objetivo”.

Para el representacionista todo acto de pensamiento involucra ideas y sólo ideas. Cualquier afirmación sólo puede ser entendida si se la escucha como haciendo referencia, en primer lugar, a las propias ideas del hablante y sólo de modo secundario, derivativo, a los hechos del mundo de los que se habla. Esto mismo puede valer para el fenomenalismo: el lenguaje solo se refiere a las “entidades internas” y de un modo derivativo (o de ningún modo) a los hechos externos. Por eso, si hacemos una interpretación representacionista – fenomenalista de la referencia primaria del lenguaje, la expresión “Yo tengo envidia de José” debe entenderse como “Tengo la idea (o la “entidad interna”) de mí mismo, más la idea (o “entidad interna”) de José y entre ambas ideas media la idea de envidia”. Esta reconstrucción resulta inaceptable, pues implicaría un desconocimiento de la naturaleza de fenómenos mentales tan cotidianos como la envidia, los celos o cualquier otro de naturaleza triádica en los que alguno de los términos de esa relación es un objeto exterior. Veíamos, sin embargo, que en el realismo por representación, el lenguaje se refiere a la idea representada y

sólo a través de ella puede referirse a un hecho externo. Por ello, los eslabones de significado son tortuosamente indirectos y falibles: Lenguaje – Idea – Mundo Exterior. Dado que podemos estar profundamente equivocados acerca de la naturaleza del mundo exterior (o, como en el idealismo de Berkeley, el mundo exterior quizás no tenga existencia), la única certidumbre que poseemos es la relación entre nuestras propias ideas, y en rigor sólo a partir de ellas puede hacerse el “salto” para afirmar un hecho del mundo. Sin referencia primaria (las ideas) no hay referencia secundaria (el mundo exterior).

Esta mediación de la idea entre el lenguaje y el mundo parece demasiado fuerte. En primer lugar, la *disunidad* de la noción de “idea” está provocando malentendidos. En segundo lugar, no necesitamos, en la mayoría de los contextos, remitirnos a una idea para referirnos a un objeto, si por “idea” entendemos “representación”.

Algunos de los malentendidos provocados por la palabra “idea” están dados por el uso nominalista que le otorgaron tanto Locke como Berkeley al término. Todas las entidades son particulares; las ideas deben serlo también, y por ello deben tener rasgos particulares. Los rasgos particulares de las ideas son los modos de darse puramente individuales, subjetivos, de dichas ideas. Es esencial, entonces, a la idea, un modo de darse idiótico, intransferible y particular. La generalidad es un rasgo del lenguaje, no de las ideas, y ciertamente es por medio del lenguaje que podemos pensar ideas generales. Pero dicha generalidad, en lo que atañe a las propias ideas, no es más que un modo de relación entre ellas, cuyo contenido es siempre particular. Dado que los hechos del mundo son también particulares, las ideas pueden captar estos hechos del mundo, mientras que el lenguaje, por su carácter general, no puede hacerlo. Este esquema en esencia no difiere del aristotélico (Rabossi, 1982), aunque con una importante salvedad: mientras la relación lenguaje – pensamiento es transparente, la relación pensamiento – mundo se ha vuelto oscura, falible e incierta.

En segundo lugar, en la mayoría de los contextos cotidianos y científicos no deseamos hablar de las relaciones entre nuestras ideas (tal como en el ejemplo de la envidia), sino de relaciones entre hechos del mundo o de relaciones entre nuestras ideas y el mundo. Pero el internismo por representación nos dirá, contra esta intención y contra

nuestra intuición de los usos del lenguaje, que en realidad estamos hablando de nuestras ideas y que sólo a través de ellas realizamos el salto hacia el mundo exterior. Para decirlo con más propiedad, no es el lenguaje el que se comunica con el mundo, sino las ideas, y en esto se sigue estrictamente el esquema aristotélico – en tanto que, en dicho esquema, el lenguaje no adquiere su referencia a partir de la realidad, sino a partir de una mediación que realiza la mente (cf.: Rabossi, 1982). El contenido comunicativo de una proposición del tipo “Yo siento envidia de José” es la transmisión de ciertas ideas, aun cuando el receptor haga una interpretación en la que la referencia es un determinado hecho del mundo. Si José escucha esa proposición pronunciada por mí, reaccionará sabiendo que él forma parte de la intención terminal de mi alocución. Sin embargo, esa reacción sólo se deriva de mis ideas, no del lenguaje; y José no tiene acceso a mis ideas. José infiere mis ideas a partir de mi lenguaje, y de ese modo adivina que mis ideas lo representan a él y representan envidia cuando lo representan a él. Dado que José asume que las ideas que lo representan a él están conectadas de algún modo con él, José se inquieta (o se alegra) por provocar esa conexión entre ideas “José + Envidia”. Todo este tortuoso camino nos conduce a una pregunta inquietante: ¿Qué ocurre si José no desanda el camino inferencial que parece necesario para entender cabalmente la proposición? Un punto importante de este camino es que el receptor se forme en su mente las mismas ideas - y la misma conexión entre ellas- que se formó el emisor. Sin embargo esto, entendido de modo literal, es imposible, pues las ideas están atadas a determinados modos de darse puramente idióticos, propios de cada individuo, y cada instanciación de una misma idea por parte de individuos diferentes provoca, *ipso facto*, que la idea no sea la misma. Las relaciones de identidad entre *lo que se dijo* y *lo que se entendió* se ven profundamente contrariadas, pues lo que yo dije sólo se refiere a mis ideas, pero lo que José entiende sólo puede referirse a las ideas de él. Por supuesto, no es necesario entender este proceso de manera literal; podemos pensar que la idea-tipo es la misma tanto en José como en mí, mientras que la idea-caso difiere, quizás sólo numéricamente. Sin embargo es cierto que el modo de darse de la idea forma parte esencial de dicha idea, y que ese modo de darse es subjetivo y que no hay manera de encontrar congruencia entre los modos de darse de José y los míos. No hay modo de

encontrar una relación de identidad entre mis ideas y las de José, aun cuando tengan el mismo contenido representacional.

Y aún hay más. ¿Qué ocurre si José no piensa una idea con contenido representacional? ¿Es necesario que José se figure a “José”, a los “celos” y a “Yo” en su mente, para que digamos que José ha entendido la proposición? Supongamos que escucha la proposición y dice haberla entendido (y reacciona como si la hubiera entendido: increpándonos, preguntándonos por qué le tenemos envidia), pero no es capaz de descubrir un conjunto triádico de ideas particulares relacionadas. Recordemos que el presupuesto básico del realismo por representación nominalista es que, para que la comprensión esté garantizada, se requiere como condición necesaria que el lenguaje esté conectado con ciertas ideas particulares, las cuales tendrán características notoriamente particulares. Parece necesario, por ejemplo, que José se haya representado a “José” desde algún ángulo, peinado de determinada manera, con la mirada serena o extraviada, sentado o de pie (o sin cuerpo, como en una fotografía tipo carnet). De algún modo, José debería ser capaz de describir su idea, y si no fuera capaz de ello, el realismo por representación podría concluir que no ha entendido la expresión.

Estas objeciones podrían llevarnos a admitir que en la mayoría de los casos no entendemos lo que se nos dice, o que transformamos el contenido de lo que se nos dice (un contenido cuya referencia sólo está en las ideas del emisor) para adaptarlo a nuestras ideas en tanto que receptores. Esta posición no es, sin embargo, incongruente con algunos aspectos de nuestra vida social cotidiana. José ama a Rebeca, pero Rebeca le objeta que José sólo se enamoró de una falsa idea de lo que es Rebeca. José ha idealizado a Rebeca, y está enamorado de esa idea, no de la auténtica Rebeca. En esta objeción de Rebeca está presente la noción de que es posible equivocarse con respecto a la propia intención terminal del lenguaje: cuando José dijo “Yo amo a Rebeca”, confundió su propia idea (la intención primaria de su expresión) con la auténtica Rebeca (la intención secundaria). La conexión entre la idea y el mundo es tan falible, que hasta que José no dé señales inequívocas de que ha comenzado a hacer congruente la idea de Rebeca con la auténtica Rebeca, es posible que Rebeca no corresponda a su amor. Nótese que esta objeción de Rebeca podría hacerse extensiva

a toda atribución triádica entre personas: quizás cuando yo dije que estaba envidioso de José, en realidad sólo tenía envidia de la idea que me formé de José. Quizás creía que José tenía mucho éxito para conquistar mujeres (y por eso le tenía envidia), pero resulta que José está enamorado de Rebeca y Rebeca no le corresponde. De modo que no podía tenerle envidia a José a causa de un hecho que no puede serle atribuido. En conclusión, yo no tenía envidia del auténtico José, sino de la falsa idea que me había hecho de él.

Ahora bien, si estamos sistemáticamente equivocados con respecto a las atribuciones de nuestras ideas, atribuciones mediante las cuales pretendemos conectar una idea con un hecho del mundo exterior, nos veríamos obligados a abandonar la intención secundaria del lenguaje. ¿Qué objeto tendría hablar de José o de Rebeca si, en definitiva, el contenido de mis ideas es sistemáticamente inadecuado con respecto al José y a la Rebeca reales? O, peor aún, aun cuando dicho contenido fuese adecuado, yo no tendría medio de corroborar si efectivamente lo es o no, y ni siquiera Rebeca o José, como intenciones terminales de mis proposiciones sobre ellos, podrían confirmarnos si nuestras ideas son correctas o no lo son, ¡pues las ideas de ellos sobre lo que yo les atribuyo difieren de las mías!³⁸

En lo que respecta a las acciones humanas siempre queda un margen para sospechar que nuestras atribuciones son erróneas, pues tales atribuciones son a su vez pensadas por esas personas, interpretadas a la luz del modo de darse en ellas el entendimiento de tales atribuciones, y ellas mismas pueden suministrarnos un juicio acerca de nuestro parecer sobre ellas, lo cual configura y reconfigura el contenido de las

³⁸ Es importante destacar que, aun cuando el lenguaje tiene una idea como referencia primaria, esto no significa que dicha idea sea una "imagen" en el sentido pictórico del término. Se puede tener como referencia una idea abstracta o una idea general. Es conocida la disputa entre Locke y Berkeley acerca de la realidad de las ideas abstractas. Para Locke, es posible tener una idea general abstracta, i.e., una idea cuya realidad objetiva no posea ninguna de las propiedades que suelen caracterizar al hecho percibido y que solo contenga la esencia (nominal) de ese hecho. En otras palabras, para Locke, es posible referirse a la "idea de silla" sin que tengamos presente una imagen pictórica de una silla. Para Berkeley tal cosa es imposible o, al menos, en cada idea debe ir *asociada* alguna cualidad específica, pues no es asequible a la mente humana una idea abstracta, sin contenido representativo. Ahora bien: es posible mentar conceptos generales, aunque no abstractos. Para Berkeley, nuestro intelecto utiliza una idea que tiene algún tipo de componente representativo como referencia de ese concepto. Así, la palabra "Hombre" podría referirse a una idea con un contenido pictórico particular (por ejemplo, un hombre determinado o un esquema pictórico de la figura humana) que sirva de signo al concepto de hombre. En rigor no es esa la "idea de hombre", sino una idea-signo (con características particulares) que representa a todos los hombres. Cf. Berkeley, (1710) Introducción, § XII.

atribuciones que les endilgamos. Decimos que Rebeca es caprichosa porque no accede al amor de José, pero quizás Rebeca no accede al amor de José porque desconfía del contenido de ese amor. De modo que ella rechazaría la atribución de “caprichosa”, en tanto sólo está exigiendo ciertas conductas para corroborar cierta atribución en José. Pero José considera que la exigencia de dichas conductas se encuadra perfectamente en la atribución de “caprichosidad”. Las personas, como objetos de nuestro pensamiento, se comportan de forma activa con respecto a nuestras atribuciones, y pueden rechazarlas, aceptarlas, refinarlas o discutir las. Pero cuando el objeto de nuestro pensamiento no está dirigido a otras personas, sino a hechos del mundo, tal ida y vuelta no es posible. La expresión “La Luna es menor que la Tierra” se refiere, según el esquema representacionista, en primera instancia a una relación entre ideas y sólo de manera secundaria designa un hecho del mundo³⁹. Ni la Luna ni la Tierra pueden comunicarnos mediante el lenguaje si tal hecho les corresponde o no, pero mediante métodos intersubjetivos es posible obtener criterios para aceptar o rechazar esta proposición. Aun cuando los objetos involucrados no pueden decirnos si las atribuciones que les realizamos son correctas, otras personas pueden confirmarnos que los métodos utilizados para corroborar esta afirmación son correctos o no lo son. Necesitamos del juicio intersubjetivo para establecer la verdad de las atribuciones. Pero, una vez más, tal juicio intersubjetivo tiene, como referencia primaria, las ideas de los sujetos que están en condiciones de emitir tal juicio. De modo que la intersubjetividad está anclada en criterios subjetivos (el carácter idiosincrásico de las ideas de cada sujeto que emite el juicio) y criterios tales son precisamente la *ausencia*

³⁹ “(...) [L]os objetos intencionales inmediatos de los estados mentales no son objetos reales ni sus propiedades, sino entidades mentales (ideas)” (García Carpintero, 1996, pág. 76)

“No es literalmente verdad que yo esté percibiendo una situación que objetivamente contiene una superficie roja. «Lo» que yo percibo, inmediatamente, son mis ideas” (Pág. 77)

“[E]l objeto intencional al que los estados mentales hacen referencia, y al que están dirigidos, no es una “realidad”; es decir: el objeto intencional de un estado mental prototípico no tiene por qué existir. Esta es la primera de las características de los estados intencionales, su *falibilidad*. Se trata de una falibilidad extrema, pues el representacionista supone que hipótesis escépticas radicales como la del Genio Maligno son lógicamente coherentes: cabe dudar de todas nuestras convicciones sobre el mundo de los acontecimientos objetivos. De lo único de que no cabe dudar es de que tenemos esas convicciones, con esos objetos intencionales. Los representacionistas precisan explicar esto postulando, para dar cuenta del contenido proposicional de uno cualquiera de estos estados, exclusivamente entidades que *no* son constituyentes de acontecimientos, sino sólo de vivencias: entidades, en suma, que sólo existen *inmanente* al estado mental. (Pág. 78-9)

de criterio. Un criterio tal es como la tabla que sólo existe en la imaginación propuesta por Wittgenstein en el § 265 de las Investigaciones Filosóficas:

Imaginémonos una tabla que existiese sólo en nuestra imaginación; algo así como un diccionario. Mediante un diccionario se puede justificar la traducción de una palabra X por una palabra Y. ¿Pero debemos también decir que se trata de una justificación cuando esta tabla sólo se consulta en la imaginación? - «Bueno, entonces es precisamente una justificación subjetiva» - Pero la justificación consiste, por cierto, en apelar a una instancia independiente. (...) Consultar una tabla en la imaginación es tan poco consultar una tabla, como la imagen del resultado de un experimento es el resultado de un experimento” (Wittgenstein, 1988, pág. 229)

Es la ausencia de una instancia independiente la que malogra la intención secundaria del lenguaje.

El idealismo y el realismo por representación parecen asumir que todo hecho mental debe tener, por un lado, contenido representacional y, por el otro, que no existen medios seguros para decidir si existe algo afuera de la propia mente. Pero, ¿es en rigor necesaria esa decisión? Podemos oponer las posturas idealista y realista por representación a lo que llamaremos *realismo intencionalista* de Frege:

El realismo intencionalista asume que la postulación de la referencia (Bedeutung) secundaria de un término incluye una *presuposición de la existencia independiente de la entidad que sirve de referencia secundaria*, aun cuando se mantenga neutral con respecto al hecho de si esa entidad existe o no efectivamente.

En *Sobre Sentido y Referencia*, (1985) Frege expone claramente esta posición:

De parte idealista o escéptica, a todo esto quizá se habrá objetado desde ya rato lo siguiente: “Hablas aquí sin más de la Luna como de un objeto. ¿Pero cómo sabes tú que el nombre «La Luna» tiene alguna referencia?; ¿Cómo sabes que hay algo que tenga referencia? Respondo que *nuestro propósito* no es hablar de nuestra representación de la Luna, y que tampoco nos conformamos con el sentido cuando decimos «La Luna», sino que *presuponemos* una referencia. Sería perder totalmente el sentido si se quisiera suponer que, en el enunciado «La Luna es menor que la tierra», se está hablando de una representación de la Luna” (Frege, 1985, pág. 85, cursivas añadidas)

El hecho de que Frege distinga entre sentido, referencia y representación contribuye decisivamente a zanjar el problema de la *disunidad* de la idea. No se trata de que el lenguaje se refiera a las ideas; se trata de que los componentes de eso que se da en llamar “idea” tienen una función diferente con respecto al lenguaje de la que suponían los modernos. Esas funciones otorgarán diferentes propiedades a la expresión, según sea su intención terminal. En expresiones como “La Luna es menor que la Tierra” la intención terminal no puede ser la idea de la Luna; la intención terminal es la referencia secundaria, esto es, la Luna, la Tierra y la determinada relación que se da entre ellas. Pero las *ideas* de Luna y de Tierra tampoco deberían ser la referencia primaria de dichos términos. Pues si aceptamos que las ideas son intermediarias entre el lenguaje y la realidad, surgen los problemas enumerados más arriba: el lenguaje está irremisiblemente preso de las ideas particulares de cada individuo, y la intención terminal se ve peligrosamente malograda. Por eso las nociones de *sentido*, *referencia* y *representación* sustituyen a la noción de *idea* de la modernidad. El punto crucial aquí es que, en el uso no intensional del lenguaje, la *referencia* (Bedeutung) nunca puede ser un hecho puramente mental como una idea. La referencia es “aquello de lo que se quiere hablar” (1985, pág. 55). En “La Luna es menor que la Tierra”, se quiere hablar de una relación externa entre objetos externos, y basta esta *intención* para establecer la salida de la propia subjetividad. No es necesario postular una idea como intermedio entre la expresión y su referencia pero, si se postulase algo así, tal idea sería lo que Frege llama *representación* (Vorstellung) y, dado su carácter puramente subjetivo, no puede formar parte relevante del lenguaje.

Los nombres propios y las expresiones funcionales poseen una referencia, lo que implica que el lenguaje está naturalmente volcado hacia los hechos externos: las expresiones no se refieren a ningún hecho mental sino a hechos externos. ¿Prueba esto, entonces, que existen esas entidades a las que se refiere el lenguaje? No lo prueba, pero no es necesario probarlo: las expresiones se comportan como si las entidades referidas existieran, fueran de dominio público y pudieran valer de manera objetiva. En cierto modo no hay lenguaje sin esta presuposición implícita. La expresión “La luna es menor que la Tierra” no es naturalmente interpretada (ni por el emisor, ni por el receptor) como: “El contenido representacional de la Luna se me figura como

menor al contenido representativo de la Tierra". Cada vez que se utilizó el término «flogisto» en la química del siglo XVIII, se pensó que se trataba de un término con referencia, y por lo tanto se asumía la existencia del objeto referido. Cuando se descubre que un término no posee referencia, se trata de un hecho científico y no de algo que pueda postularse a priori mediante un argumento escéptico o idealista⁴⁰.

El realista por representación establece un nexo sumamente débil entre el pensamiento y el mundo. El realista intencionalista no requiere de dicho nexo. Y dado que el pensamiento es estrictamente proposicional y objetivo, la representación (Vorstellung) a él asociada no puede formar parte de dicho pensamiento:

De la referencia y del sentido de un signo hay que distinguir la representación a él asociada. Si la referencia de un signo es un objeto sensiblemente perceptible, la representación que yo tengo de él es entonces una imagen interna formada a partir de recuerdos de impresiones sensibles que he tenido y de actividades que he practicado, tanto internas como externas (1985, pág. 56)

La representación está atada a un portador y a un momento, pero no el sentido:

Mientras que (...) no existe ninguna objeción para hablar del sentido sin más, en el caso de la representación en cambio, para ser estrictos, hay que añadir a quién pertenece y en qué momento. Quizás alguien diría: al igual que con la misma palabra uno asocia tal representación, el otro tal otra, también puede uno asociarle tal sentido, el otro tal otro. Sin embargo, la diferencia consiste entonces en el modo de darse esta asociación. Esto no impide que ambos conciban el mismo sentido; pero no pueden tener la misma representación. *Si duo idem faciunt, non est idem*. Cuando dos personas se representan lo mismo, cada una tiene, sin embargo, su representación propia. (1985, págs. 56-7)

Mientras el sentido es el modo intersubjetivo de darse la referencia, la representación es un modo de darse subjetivo e intransferible. ¿Pero de qué está hecha una representación? Una nota al pie parece aclararlo:

⁴⁰ Exceptuamos, desde luego, los nombres propios que han sido creados con el propósito de no tener referencia, como Pegaso o Sherlock Holmes. En esos casos, la no-referencia del término no se descubre, sino que se evidencia a priori en el contexto enunciativo. (Cf. Strawson, 1995, pág. 68: "Si empezase diciendo «El rey de Francia es sabio», y continuase «y vive en un castillo de oro y tiene cien esposas», etc., un oyente me entendería perfectamente bien, sin suponer *o bien* que estaba hablando acerca de una persona particular, *o* que estaba haciendo un enunciado falso en el sentido de que existía una persona tal como la descrita con mis palabras")

Junto a las representaciones, podemos poner también las intuiciones o los datos sensoriales, en los que las impresiones sensibles y las actividades mismas ocupan el lugar que han dejado en el espíritu. Para nuestro propósito, la diferencia es irrelevante, tanto más cuanto que, junto a las sensaciones y actividades, los recuerdos de éstas ayudan a completar la imagen intuitiva. Por intuición o dato sensorial, sin embargo, puede entenderse también un objeto, en la medida en que éste sea sensiblemente perceptible o espacial. (1985, pág. 56n)

Las representaciones son, entonces, las huellas que han dejado las cualidades, más las operaciones que la propia mente ha realizado con ellas. Sin embargo, las representaciones espaciales (esto es, las derivadas de cualidades primarias) suministran el conocimiento de un objeto. Aunque tienen un componente representativo, dicho componente sólo es relevante en tanto permite el establecimiento de una entidad intersubjetiva con propiedades geométricas. Esta afirmación es totalmente compatible con la tesis de la convergencia geométrica en Locke (cf. § 1.5.2) y con los requisitos C-C-C (cf. § 1.2.1) de las cualidades primarias. Sin embargo, a diferencia de Locke, Frege no considera necesario justificar esta relación entre las representaciones espaciales y el objeto, pues la noción de *sentido* funciona como elemento nómico validador del carácter objetivo de esa relación “representación – objeto espacial”. El sentido, como modo de darse la referencia, es un componente *intersubjetivo* de los distintos modos en los que una referencia puede darse a un individuo a través del lenguaje. Dado que las propiedades geométricas cumplen con este carácter de intersubjetividad, comparten algunas características del *sentido*. De hecho, la única manera de aceptar que las representaciones espaciales suministran un objeto estriba en entender a dicha representación espacial como portadora de un sentido, “en la medida en que puede servir a varios observadores” (Frege, 1985, pág. 57)

Los argumentos contra el escéptico y el idealista, brevemente esbozados en *Sobre Sentido y Referencia*, tendrán un mayor desarrollo en *El Pensamiento*. Dado que un pensamiento es el sentido de una proposición, es necesario separar de la noción de “pensamiento” todo elemento subjetivo. ¿Por qué hay que aceptar que los pensamientos difieren de las representaciones? Locke, Descartes y Berkeley utilizaban el término *idea* para referirse a cualquier contenido mental. Las representaciones requieren de un portador y de un tiempo. Si el pensamiento se identifica con una

representación, cabe la posibilidad de preguntarse si un mismo pensamiento puede ser pronunciado por dos personas: ¿cómo puede establecerse una identidad entre nociones que requieren de estados mentales diferentes para su instanciación?:

¿Es en efecto el mismo pensamiento, el que pronuncia primero aquella persona y ahora ésta? El hombre no tocado aun por la filosofía conoce inmediatamente cosas que puede ver, tocar, en suma, percibir con los sentidos, tales como árboles, piedras, casas; y está convencido de que otro hombre puede ver y tocar del mismo modo el mismo árbol. Es evidente que un pensamiento no pertenece a esta clase de cosas. ¿Pero puede él, a pesar de ello, presentarse a los hombres como él mismo, igual que un árbol?

También el hombre no filosófico se ve obligado a reconocer un mundo exterior, un mundo de las impresiones sensibles, de las creaciones de su imaginación, de las sensaciones, de los sentimientos y estados de ánimo, un mundo de las inclinaciones, deseos y decisiones. Para acuñar una breve expresión, resumiré todo esto – exceptuando las decisiones- con la palabra “representación” (Frege, *El Pensamiento, Una Investigación Lógica*, 1974, pág. 174)

El reconocimiento del mundo exterior es algo propio de quienes no tienen pretensiones filosóficas. Pero la afirmación de que existe un mundo interior sólo es posible si se acepta el supuesto de que el sujeto es portador de ese mundo: “El mundo interior tiene como supuesto a uno, del que él es mundo interior” (El pensamiento, 145). Por ello, Frege enumera las cuatro propiedades de segundo orden de las representaciones:

1. Las representaciones no pueden ser vistas o tocadas, ni olidas, ni gustadas ni oídas.
2. Las representaciones se tienen: pertenecen al contenido de la conciencia.
3. Necesitan de un portador.
4. Cada representación sólo puede poseer un solo portador. Dos personas no tienen la misma representación. (1974, págs. 145-6)

Pero, ¿es posible comparar dos representaciones de dos sujetos diferentes? ¿Qué ocurre si se desea preguntar por la identidad de contenido de dos representaciones?:

Mi acompañante y yo estamos convencidos de que los dos vemos el mismo prado; pero cada uno de nosotros tiene una particular impresión sensible de lo verde. Diviso una fresa entre las hojas verdes. Mi acompañante no la encuentra; es daltoniano. La impresión de color que él recibe de la fresa no se diferencia notablemente de la que recibe de sus hojas. ¿Ve mi acompañante roja la hoja verde o ve verde la fresa roja? ¿O

ve ambas de un color que yo no conozco? Son preguntas sin respuesta, o más propiamente, preguntas sin sentido. Pues la palabra «rojo», si es que no da cuenta de una cualidad de los objetos, sino caracteriza impresiones sensibles pertenecientes a mi conciencia, es, entonces, aplicable sólo en el campo de mi conciencia, ya que es imposible comparar mi impresión sensible con la de otro. (1974, pág. 146)

Este es el problema de los *qualia invertidos*, que ya había aparecido en Locke (cfr. § 1.6) y había sido desechado por improbable, aunque no generaba confusión mientras los hablantes no confundieran cada uno su idea con la del otro. Desde luego, esta argumentación no parece sostenible: en rigor no existe la posibilidad de que se confundan las ideas, porque la idea ajena nunca está dada, de modo que yo no podría confundir mi idea de violeta con la idea de violeta de mi acompañante. Lo que Locke quería decir es que, a pesar de que las ideas que sirven de referencia primaria son diferentes en ambos casos (en número y quizás en cualidad), la comunicación descansa sobre la identidad causal de la cualidad que provoca esa idea. Y hemos visto en párrafos anteriores que es precisamente el propio marco causal el que resulta problemático. Frege también rechaza la posibilidad de los *qualia invertidos*, pero su escueto argumento es que tal planteo carece de sentido. ¿Por qué carece de sentido? Porque los modos de darse privados son inconmensurables. Si el espacio cromático de A difiere del espacio cromático de B, este es un hecho anterior a toda manifestación lingüística, y por ese motivo no puede llegar a plantearse. Por el contrario, si hay lenguaje, hay pensamiento y si hay pensamiento se ha llegado a establecer un suelo común de sentido sobre el cual hablar. Preguntar por algo que es lógicamente anterior al sentido es, precisamente, preguntar por algo que no tiene sentido. Esta aseveración convierte en *inefables* los modos de darse de la representación.

Volvamos al problema inicial, acerca de si son representaciones los pensamientos:

Si el pensamiento que pronuncio en el teorema de Pitágoras puede ser reconocido como verdadero tanto por otros como por mí, no pertenece, entonces, al contenido de mi conciencia, no soy yo, por consiguiente, su portador y, sin embargo, puedo reconocerlo como verdadero (...)

Si cada pensamiento necesita de un portador, a cuyos contenidos de conciencia pertenece, entonces sólo es pensamiento de ese portador, y no hay una ciencia común a muchos en las que muchos puedan trabajar, sino que yo tengo tal vez mi ciencia, es decir, un conjunto de pensamientos de los que soy portador y otro tiene su ciencia. Cada uno se ocupa de contenidos de su conciencia. (...) Así pues, el resultado

parece ser: los pensamientos no son ni objetos del mundo exterior ni representaciones.

Hay que considerar una tercera esfera. (Frege, 1974, págs. 147-8)

Las empresas científicas son, por antonomasia, empresas colectivas que descubren (o quizás acuñan) entidades en una tercera esfera. Se puede pensar, entonces, que si el trabajo mancomunado de los científicos hace progresar a una ciencia, ello se debe a que pueden superar las instancias puramente subjetivas de las representaciones para ponerse en contacto con un tesoro común, el de los pensamientos relativos a esa ciencia. Pero este condicional podría ser falso, porque no se ha mostrado todavía que la ciencia pueda progresar gracias al contacto con los pensamientos. Es, en todo caso, una cuestión de hecho que la ciencia avanza, pero de este hecho no puede deducirse, sin más, que progresa porque trabaja con entidades de un tercer reino. Por eso, si bien el condicional parece sumamente plausible, el trabajo de probar la existencia de un reino objetivo todavía queda por hacerse.

Para ello Frege va a realizar un recorrido inferencial que recuerda mucho a las *meditaciones metafísicas* cartesianas:

Pero me parece oír una curiosa objeción. He admitido varias veces que el mismo objeto que yo veo pueda ser contemplado también por otro. ¿Pero cómo, si todo fuera sólo un sueño? (...) quizás el reino de los objetos está vacío, y no veo objetos, tampoco seres humanos, sino tengo sólo tal vez representaciones cuyo portador soy yo mismo. (1974, pág. 148)

Si sólo hay representaciones, entonces los verbos “ver, oír, oler, tocar, gustar” no tienen referencia, pues estos verbos expresan el éxito de la empresa perceptiva. Sólo puede verse un objeto. Pero si todo es representación, entonces en rigor no vemos. Quizás haya cosas afuera, pero nuestro acceso estaría completamente vedado. La visión viene implicada por el funcionamiento de la estructura perceptiva, de la cual el nervio óptico y la retina son una parte. Pero los términos “nervio óptico” y “retina” o bien se refieren a algo que es una representación mía, o bien a algo que está fuera de ella. Como no tengo contacto con lo que está fuera de mis representaciones, entonces no tiene sentido hablar de retinas o nervios ópticos:

En rigor esta excitación del nervio óptico no se da de manera inmediata, eso es sólo una suposición. Creemos que un objeto independiente de nosotros excita un nervio y mediante ello produce una impresión sensible, pero en realidad experimentamos de este proceso sólo el último paso, que penetra en nuestra conciencia. (...) Si lo que pertenece a nuestra conciencia llamamos representación, en rigor experimentamos sólo representaciones pero no sus causas. (1974, pág. 150)

Sin embargo, parece existir algo que no es una representación:

Me he considerado portador de mis representaciones, pero ¿no soy yo mismo una representación? (...) Me parece ver también allí una silla. Es una representación. En rigor yo no difiero mucho de ella, pues acaso ¿no soy yo también una unión de impresiones sensibles, una representación? ¿Cómo he llegado a escoger una de esas representaciones y a proponerla como portadora de las otras? ¿Por qué ha de ser esa la representación que he dado en llamar *yo*? ¿No podría elegir para ello del mismo modo la representación que estoy tratando de llamar silla? (...) Si todo es representación, entonces no hay ningún portador de las representaciones. (...) ¿Pero es posible eso? ¿Puede haber una experiencia sin alguien que la experimente? ¿Qué sería de este gran espectáculo sin un espectador? (1974, pág. 151)

Si todo es representación, la propia noción de *portador* es arbitraria. Pero la propia noción de representación se desvanece si no se presupone al portador. De ese modo, se prueba que el portador debe ser algo diferente, una condición de posibilidad de las experiencias, pero no puede ser a su vez una experiencia más:

Tengo una representación de mí, pero yo no soy esa representación. Hay que diferenciar rigurosamente aquello que es contenido de mi conciencia, la representación, y aquello que es objeto de mi pensamiento. Así pues, es falsa la proposición de que sólo puede ser objeto de mi contemplación, de mi pensamiento, lo que pertenece al contenido de mi conciencia. (1974, pág. 153)

Se ha dado un paso decisivo: la afirmación de que no todo *lo que pasa* en mi mente pertenece a ella. Ninguna afirmación de este tipo podía encontrarse en el pensamiento cartesiano, en el cual existe una identidad entre una idea (definida como cualquier evento consciente) y un contenido de conciencia. Se pueden distinguir los *contenidos de conciencia* de los *objetos de pensamiento* o, dicho de otro modo, el acto de pensar es esencialmente diferente del acto representativo. A los objetos de pensamiento se

los *aprehende*, pero no se los tiene como parte de la conciencia. En una nota al pie, Frege muestra una analogía para comprender la distinción entre el *aprehender* y *formar parte*:

La expresión “aprehender” es tan ilustrativa como “contenido de conciencia”. La esencia del lenguaje, justamente, no lo podría permitir de otra manera. Lo que tengo en mi mano puede ser considerado contenido de ella, pero es contenido de la mano de manera completamente distinta y mucho más extraño a ella que los huesos, que los músculos de que ella se compone y que sus tensiones. (1974, pág. 153n)

Las representaciones en la conciencia son como los huesos de la mano: son parte integrante de ella. La conciencia puede, sin embargo, asir algo que no le pertenece de modo esencial. Y *sólo puede asirse algo si previamente ese algo ya existía de manera independiente*: “La tarea de la ciencia no consiste en un crear, sino en un descubrir pensamientos verdaderos” (1974, pág. 154) Por eso, hay que admitir la existencia de un tercer reino, un reino no físico y no mental, pero absolutamente real. Los pensamientos son captados por una facultad especial, que nada tiene que ver con la capacidad de representación: el intelecto.

Una vez aceptado el hecho de que no todo es un contenido de conciencia, se puede aceptar que existan otras personas que sean portadoras independientes de representaciones (1974, pág. 152). El camino está libre para admitir que lo que es real puede ser aprehendido, aun cuando no sea un contenido de conciencia. Los hechos son objeto de pensamiento, aunque ellos mismos no sean contenido de conciencia: “¿Qué es un hecho? Un hecho es un pensamiento que es verdadero” (1974, pág. 154). Así, el conocimiento de la realidad de los hechos del mundo empírico no se deriva de las sensaciones, sino de la aprehensión de la verdad de un pensamiento. La conciencia debe “apuntar” al pensamiento para aprehenderlo, pero la operación de la conciencia no debe ser confundida con el pensamiento mismo.

Los pensamientos son inactuales, no cambiantes y no pertenecientes al mundo interior de los individuos: pertenecen a un tercer reino. “¿Pero no hay también pensamientos que son verdaderos, pero después de medio año falsos? Por ejemplo, ¿el pensamiento de que aquel árbol tiene el follaje verde, es, después de medio año, falso? No, porque no es el mismo pensamiento” (1974, pág. 155). Las expresiones indexicales (aquel,

éste, hoy) pertenecen a la expresión del pensamiento, pero no son parte del pensamiento. Lo que cambian son las circunstancias en las que es posible decir un pensamiento verdadero, pero el pensamiento en sí mismo no cambia. “El presente de «es verdadero» no se refiere al momento en que es dicho, sino es, si se me permite la expresión, un tiempo de la atemporalidad” (1974, pág. 156) Ahora bien, ¿cómo puede actuar en nosotros lo que es inactual? ¿Cómo entra un pensamiento en contacto con nuestra conciencia? A través de sus cualidades inesenciales: el “ser aprehendido hoy” es una cualidad inesencial del pensamiento. “¿Cómo actúa un pensamiento? Siendo aprehendido y tenido por verdadero. Es un proceso en el mundo interior de un [sujeto] pensante y que puede tener posteriores consecuencias en ese mundo interior” (1974, pág. 156)

La apelación a las cualidades inesenciales es un expediente poco claro: el pensamiento actúa siendo aprehendido, y su “ser aprehendido” es la cualidad no esencial gracias a la cual entramos en contacto con él. Nótese que su actuación está en voz pasiva: actúa siendo reconocido como verdadero e inactual: “Los pensamientos no son del todo inactuales, pero su actualidad es de otra índole que la de los objetos. Y su actuar es desencadenado por un hacer del sujeto pensante, sin el cual –al menos en la medida en que nos es posible verlo- serían inefectivos” (1974, pág. 157) Resulta curioso, además, que sólo entremos en contacto por medio de cualidades inesenciales. La posibilidad de ser captado es una de esas cualidades inesenciales. En tanto pensamos que el pensamiento cambia, estaremos en el terreno de la cualidad inesencial. Pero, entonces, la mayoría de las personas, en tanto capte pensamientos cuya verdad está sujeta a circunstancias temporales, sólo capta cualidades inesenciales; sólo capta la posibilidad de un pensamiento y no un pensamiento. Podemos pensar al pensamiento como una entidad objetiva con dos núcleos: el núcleo duro, central, formado por sus cualidades esenciales: atemporalidad, inactualidad, objetividad. El núcleo periférico, que rodea al núcleo central, está formado por sus cualidades inesenciales: posibilidad de ser captado, temporalidad de la captación, cambio de valor de verdad con el paso del tiempo. Frege busca una salida del realismo por representación: no todo lo que pasa por nuestra conciencia es representación, ni pertenece a ella. Sin embargo, esta salida tiene sus propios problemas, pues la noción de “captación”, junto con la

atemporalidad e inactualidad de lo captado, necesitan de aclaraciones ulteriores: se requiere de una psicología de esta peculiar "captación", y de una caracterización satisfactoria que pueda dar cuenta de la captación entre entidades inconmensurables: la conciencia humana, temporal y actual, y los pensamientos, inactuales e intemporales. Para ponerlo en palabras de Hilary Putnam (1995), a propósito del pensamiento y del sentido entendidos como significado o intensión de los términos:

Muchos filósofos tradicionales concibieron el concepto como algo mental. Así, la doctrina de que el significado de un término (es decir, «en el sentido de intensión») es un concepto que llevaba consigo la consecuencia de que los significados son entidades mentales. Frege, y más recientemente Carnap y sus seguidores se han rebelado, no obstante, contra este psicologismo. Apercibiéndose de que los significados son propiedad *pública* –que el *mismo* significado puede ser «captado» por más de una persona y por personas situadas en tiempos diferentes- sostuvieron que los conceptos (y, por lo tanto, las «intensiones» o significados) eran entidades abstractas, en vez de entidades mentales. Sin embargo, «captar» estas entidades abstractas siguió siendo un acto psicológico individual. Ninguno de estos filósofos puso en duda que entender una palabra (conocer su intensión) no fuese precisamente cosa de estar en un cierto estado psicológico (de alguna forma, al modo en que saber cómo extraer mentalmente los factores de un número es exactamente cosa de estar en un cierto estado psicológico muy complejo) (Putnam, El significado de "Significado", 1995, pág. 135)

En su lucha contra el realismo por representación, Frege queda preso de una noción no menos problemática que la de la idea moderna: la noción de captación como acto psicológico individual que, sin embargo, tiene características idénticas para todo individuo, pues lo captado habrá de ser siempre objetivo. Tendremos presente, de todos modos, que su aporte ofrece una propuesta correctiva que será retomada al final de este trabajo: la noción de que no todo lo que es contenido de conciencia deba ser representación. Hay espacio, entonces, para que podamos aprehender algo que no es una pura representación, sino que es, en sí mismo, un aspecto de la realidad objetiva.

1.9. Balance: el realismo por representación versus el realismo intencionalista.

Locke no dio con una teoría satisfactoria para explicar la relación entre los estados de conciencia y la realidad exterior. La tesis mecanicista corpuscular, aun siendo “la hipótesis más racional” (§§ 1.5, 1.7), no podía dar cuenta de las ideas provenientes de cualidades secundarias, razón por la cual era necesario apelar a la arbitrariedad divina para explicar la aparición de un color *qua* color en la conciencia. Dado que los contenidos inmediatos de conciencia son todo lo que se dispone, las ideas –tanto las que son provenientes de cualidades primarias como de cualidades secundarias- sirven como puntos de partida para inferir la existencia de cualidades y de esencias reales. Pero las ideas provenientes de cualidades secundarias no pueden utilizarse como puntapiés para inferir la existencia de un mundo exterior, pues no hay semejanza entre dichas ideas y el hecho externo que les da origen. Dado que todos los hechos mentales poseen un contenido representacional en principio, y dado que este contenido representacional necesita de características particulares, toda representación contiene ideas que remiten a cualidades secundarias. La consecuencia de esto es que en cualquier acto de conciencia están presentes ideas que no tienen semejanza con el mundo exterior. Incluso en aquellas ideas en las que supuestamente sí hay una semejanza, tal semejanza no puede ser con la realidad última, esto es, con las “cualidades reales”. Dado que la conexión entre los fenómenos y los corpúsculos es postulada *de facto*, aunque no *de iure*, los caminos hacia el idealismo, el escepticismo o el fenomenalismo están abiertos. En cualquiera de las tres vertientes se acepta que no es posible el contacto con la realidad exterior. Esta afirmación comporta no sólo el fracaso de las teorías científicas, sino también el sinsentido de todos los enunciados de nuestra vida cotidiana acerca de relaciones entre hechos del mundo: todo hecho es en rigor un hecho mental, y toda relación es una relación en una mente.

El realismo intencionalista de Frege se libera de la teoría representacionalista. Las representaciones no tienen lugar ni en la ciencia, ni como parte importante de la referencia de las expresiones lingüísticas. Por lo tanto, las afirmaciones acerca de hechos del mundo no requieren de un determinado hecho mental con características representacionales. La mente puede captar pensamientos, pero los pensamientos

poseen una naturaleza muy distinta a la de las representaciones. Mientras una representación se tiene, el pensamiento se capta. La captación ya presupone la realidad de lo captado. El realismo intencional de Frege no debe confundirse con el “realismo fingido”. Según García Carpintero:

La corrección [con respecto al realismo por representación y al fenomenalismo] que caracteriza al *realismo fingido* sobre un ámbito del discurso consiste en la tesis de que las entidades características de ese discurso son entidades ficticias: hablamos y pensamos *como si* existieran realmente (y no es necesario ni conveniente dejar de hablar así: en esto radica la diferencia con el reductivista eliminador), pero a todos los efectos intelectuales y prácticos, su estatuto ontológico es el de las entidades de la ficción. Esto no significa que el realismo fingido las considere “ficticias”, sino que se trata de entidades que no pueden tener incidencia alguna en la justificación de lo que decimos o de lo que pensamos, en la determinación de si pensamos y hablamos correctamente sobre ellos o no lo hacemos. (1996, pág. 157)

El realismo fingido postula que nunca estamos en condiciones de conocer la verdad sobre un determinado tema. Es una doctrina para la cual las proposiciones y los términos teóricos sólo tienen un valor instrumental. Sin embargo las diferencias con el realismo intencional son notorias. Frege afirma la existencia de las numerosas entidades del tercer mundo que sirven de referencia a las expresiones: las entidades matemáticas, lo verdadero, lo falso, las funciones, los conceptos y los pensamientos (Frege, *Función y concepto*, 1974). La captación intelectual sólo se da en el terreno de este tercer reino, que es objetivo y captable de manera idéntica por todo sujeto que posea intelecto. Las sensaciones, en cambio, nos otorgan sólo representaciones, aunque algunas de éstas – las que poseen contenido espacial- suministran un objeto: no un objeto intelectual (como podrían serlo la captación de lo verdadero o lo falso), sino un objeto empírico. De qué modo el objeto empírico condiciona la verdad o la falsedad de una proposición, es algo de lo cual Frege no va a dar mayores detalles, aunque sin duda la experiencia juega algún rol en la apreciación de la verdad o la falsedad de un pensamiento. En este punto es ilustrativo el § 486 de las *Investigaciones Filosóficas* de Wittgenstein:

¿Se sigue de las impresiones sensoriales que tengo, que ahí hay un sillón? - ¿Cómo puede seguirse de las impresiones sensoriales una *proposición*? Bueno, ¿se sigue de las proposiciones que describen las impresiones sensoriales? No - ¿Pero acaso no infiero de las impresiones, de los datos sensoriales, que allí hay un sillón? - ¡No hago ninguna inferencia! – Aunque a veces sí. Veo, por ejemplo, una fotografía y digo: «Así, pues, allí

tuvo que haber un sillón», o también: «por lo que se ve ahí, infiero que allí hay un sillón». Esto es una inferencia; pero no una inferencia de la lógica. Una inferencia es la transición de una aserción; por lo tanto, también a la conducta que corresponde a la aserción. ‘Saco las consecuencias’ no sólo de palabra, sino también con acciones. (Wittgenstein, 1988, pág. 327)

¿Cuál es el lugar de los estados cualitativos de conciencia en la teoría de Frege? Debemos remarcar las diferencias y similitudes entre las nociones de *sentido* y *representación*. El sentido es el “modo de darse de lo designado” (Frege, *Sobre Sentido y Referencia*, 1985, pág. 52) y “lo comprende todo aquel que conoce el lenguaje o el conjunto de designaciones al que pertenece” (pág. 53) y “puede ser propiedad común de muchos” (pág. 56). Aunque Frege no da una definición de sentido, sus características están explícitas: el sentido de un signo es algo que puede captarse con el intelecto; es comunicable e intersubjetivo. La representación, en cambio, está confinada a la estricta individualidad. Pero, con todo, ¿no es la representación también un *modo de darse lo designado*? Es un modo, pero no el modo que los demás pueden compartir inteligiblemente. Los estados cualitativos de la conciencia (estados de “ver una superficie roja”, de “tener dolor”) son puramente privados y previos a toda actividad del intelecto. Recordemos que en Frege la facultad intelectual debe considerarse separada de la facultad representativa. Los problemas acerca de si el rojo que yo percibo es cualitativamente idéntico al rojo que percibe otra persona no son problemas genuinos, en tanto sólo tratan del contenido representativo para el cual la noción intersubjetiva de sentido no está disponible. Si se pudiera plantear de manera inteligible el problema de los qualia invertidos, eso significaría que las propias sensaciones nos transmitirían sentidos, de modo que cualquiera podría entender *lo que es que yo tenga la sensación de rojo o cómo es ser murciélago*. Esos modos de darse atados a la indexicalidad de la propia conciencia no conforman propiamente un sentido; son más bien un conjunto apiñado de fenómenos subjetivos encerrados y silenciados en el aquí y ahora de un determinado yo. Aunque Frege encuentra una salida prometedora cuando asume que puede captarse algo que no es representación, de todos modos no es la propia subjetividad la que se aclara mediante esta captación. La propia conciencia, el portador de las representaciones y sujeto de las captaciones intelectuales, permanece inexplicado. En definitiva, ni el realismo por representación,

ni el realismo intencionalista pueden saltar la brecha que media entre los estados cualitativos de conciencia y la realidad objetiva.

Segunda parte: Pareceres, qualia y la reducción de los estados cualitativos durante fines del siglo XX y principios del XXI.

El programa científico del siglo XVII realizó un intento de reducción de cualidades secundarias a primarias. Todos los cambios en los cuerpos consisten en la acción por contacto (inercia, presión y choque), puramente mecánica, de los corpúsculos de los que consta la materia. En el programa filosófico realista más extenso y fructífero del siglo XVII los estados cualitativos de conciencia no pueden explicarse de modo reductivo sin apelar a la intervención de una divinidad que, arbitrariamente, une los diversos tipos de cualidades con ideas. La propuesta de Locke de distinguir entre ideas, cualidades y esencias es difusa; la esencia real es incognoscible; las cualidades primarias de los cuerpos macroscópicos deben distinguirse de las cualidades primarias de los corpúsculos; la semejanza entre ideas y cualidades primarias sólo se puede sostener como normativa del conocimiento si tal semejanza se mantiene al nivel de las cualidades microscópicas, y sin embargo estas últimas no pueden conocerse. No hay una clara posibilidad de reducir una idea a sus causas externas: el carácter fenoménico de las ideas provenientes de cualidades secundarias no puede explicarse como superveniente a las características de las cualidades primarias sobre las que se basan las cualidades secundarias que les dan origen. A su vez, el carácter fenoménico de las ideas provenientes de cualidades primarias tampoco puede reducirse a sus causas últimas (las cualidades primarias microscópicas, propias de los corpúsculos últimos de la materia), sino tan solo pueden mantener una semejanza literal (aunque no estricta) con cualidades primarias macroscópicas. Pero estas cualidades primarias no son “cualidades reales”, sino que son reducibles a ellas. En resumen, tanto las ideas Q1 como las ideas Q2 contienen elementos fenomenológicos que no pueden explicarse reductivamente a partir de las cualidades que les dan origen.

La noción de idea es polisémica y se refiere a todo lo que es objeto de conciencia, pero a su vez hay una disunidad con respecto a este “objeto de conciencia”, pues algunas ideas contienen una señal de dirección causal, causa que no puede ser, estrictamente hablando, objeto de conciencia.

Es importante dejar asentado que ninguna de las nociones propuestas por Locke cumple satisfactoriamente con el rol conceptual que se le pretende asignar: ni la noción de idea puede caracterizar el contenido fenomenológico (debido a la introducción de un elemento causal no fenoménico), ni la noción de cualidad puede

tender el puente entre el aspecto fenomenológico y la realidad (debido a las ya señaladas fallas en la inter-reducción de ideas a cualidades, y estas a cualidades últimas, “reales”) ni, finalmente, la noción de esencia real puede servir como objeto de investigación, debido a que, a causa de la inaccesibilidad epistémica (empírica e intelectual) que supone, no es posible utilizarla como base reductiva para explicar los tipos de cualidades que se manifiestan macroscópicamente. La esencia real es una hipótesis última desconectada de cualquier intento de sistematización nómica del mundo. El marco cartesiano – las ideas del espíritu como base del conocimiento- que guía el espíritu del pensamiento de Locke es incompatible con el realismo que pretende sostener: su esquema filosófico está muy cercano al solipsismo; sin embargo se afirma la existencia del mundo externo aunque el propio esquema no puede sostener conceptualmente esta afirmación⁴¹.

En el marco del realismo intencionalista de Frege hay un intento correctivo para estos problemas: la noción de idea separa sus dos componentes disunívocos: el contenido de conciencia puramente subjetivo y el contenido de conciencia objetivo (Sinn). Sin embargo, el intento de Frege se compromete con la captación de entidades objetivas de un tercer reino, lo cual es claramente problemático: por un lado, por la necesidad de postulación de un tercer reino al estilo platónico –y las desventajas largamente señaladas del platonismo: multiplicación innecesaria de entidades, correspondencia de tales entidades con los objetos del mundo-; por el otro, por el insatisfactorio carácter de la noción de *captación* para las entidades de ese mundo: la captación es un concepto psicológico, para el cual es necesario otorgar criterios que no terminen anclados en las peculiares características psicológicas de una mente determinada –lo que haría difusa la distinción entre *sentido* y *representación*. Frege no logra dar un criterio claro y su sistema, si bien contiene nociones correctivas importantes, no puede

⁴¹ Omitimos la solución kantiana o, en rigor, las soluciones del fenomenalismo en general, en virtud de que, si bien da una respuesta a los problemas del realismo por representación (en el sentido de que los contenidos mentales no son “ideas” que “representan” a los objetos, sino que son construcciones lógicas elaboradas por la mente), no puede explicar la relación entre “lo que está afuera” (la cosa en sí) y las construcciones mentales. El idealismo trascendental kantiano no puede explicar la relación causal entre “lo interno” y “lo externo”, porque la misma relación causal es una categoría puramente “interna”. Una solución idealista no requiere de una búsqueda de puentes entre lo interno y lo externo, porque declara que lo externo o bien no tiene existencia, o bien es incognoscible. La solución aportada por el fenomenalismo kantiano no sirve a nuestro propósito, que consiste en confrontar y resaltar la continuidad de las soluciones del realismo por representación moderno con las actuales, como veremos en esta segunda parte.

explicar la brecha entre las circunstancias temporales de la captación y la inactualidad e intemporalidad de los pensamientos.

Los intentos de explicar la brecha entre el modo en que nos parecen las cosas –los estados cualitativos de conciencia, la fenomenología- y el mundo externo han sido causa de grandes desvelos (no sólo filosóficos) durante todo el siglo XX y sus resultados han sido y siguen siendo discutibles. Levine (1983) populariza la expresión de “gap” (brecha) explicativo, y Chalmers (1999) lo llama “el problema duro de la conciencia”. Según Chalmers, este “problema duro” nos muestra que las cualidades intrínsecas de la conciencia no son explicables mediante la estrategia reductiva. La dificultad para reducir los aspectos cualitativos de la conciencia a las características físicas o funcionales de los cuerpos o sistemas no se debe a que la brecha entre unos y otros requiere de hipótesis reductivas aun no formuladas pero eventualmente formulables mediante el avance y afinamiento de la investigación empírica: la dificultad estriba en que los fenómenos mentales poseen leyes propias, irreductibles y fundamentales, y por lo tanto cualquier intento de reducción estará buscando una explicación por un camino equivocado. A pesar de este diagnóstico negativo, algunos filósofos argumentan que no existe tal brecha explicativa (Daniel Dennett, entre ellos.) Otros, como Jeff Foss, dirán que el puente entre las brechas ya ha sido tendido y que la sensación de misterio que rodea al problema es producto de una mala concepción de lo que debe esperarse de una investigación científica. Otros (llamados usualmente “misterianos”) argumentarán que el conocimiento de las relaciones entre mente y mundo nos está vedado y que no tiene sentido la búsqueda de una explicación satisfactoria. Sin negar los problemas que surgen de afirmar la irreducibilidad de un dominio sobre otro (como Chalmers lo hace), es posible que quienes eliminan el problema (como Dennett) con respecto a la explicación reductiva de los fenómenos mentales a procesos funcionales o físicos estén eliminando *más* de lo que sería deseable, y que quienes suponen que la explicación *ya está dada* (como Jeff Foss), expliquen *menos* de lo deseable o sobre bases no tan sólidas como suponen.⁴²

⁴² Existen soluciones en las cuales el propio planteamiento del problema (el gap o brecha explicativa) es rechazado. En estos planteamientos, no se discute si la brecha es superable o no lo es, sino que la propia brecha entre pensamiento y mundo es una categoría inaceptable. Para la fenomenología husserliana, por ejemplo, solo se puede encontrar tal brecha en lo que Husserl denomina «la tesis de la actitud

Existen dos posiciones fundamentales con respecto a la capacidad de los estados cualitativos de conciencia de ser explicados dentro de la estructura científica naturalista: a) o bien la ciencia puede explicar, con el arsenal conceptual *actual*, todo lo que es *importante* con respecto a los estados fenoménicos; b) o bien la ciencia necesita adquirir un tipo de perspectiva diferente para capturar la índole específica del problema de los estados fenoménicos y de la conciencia en general. Dejemos de lado a), porque desde la perspectiva de quien sostiene a) no hay problema para explicar (aunque veremos que, en realidad, esa perspectiva podría basarse en asunciones erróneas) Con respecto a b), puede pensarse que: b1) o bien la ciencia jamás podrá adquirir tal perspectiva (una postura *misterianista*); b2) o bien la ciencia deberá crear un específico modo de acceso empírico y conceptual para capturar los fenómenos cualitativos de la conciencia.

Dejemos de lado b1). Si la ciencia no puede explicar los fenómenos cualitativos de la conciencia, por tratarse de hechos esencialmente alejados de cualquier investigación empírica, entonces la discusión termina aquí. B1 es, sencillamente, el dictamen a priori de que la investigación empírica (no solo actual, sino también futura) jamás podrá entender hechos cualitativos de la conciencia. Precisamente por lo arbitrario y apriorístico de este dictamen, podemos dejarlo de lado. Implica presuponer que las estructuras conceptuales *actuales* de la ciencia natural son suficientes para suponer la imposibilidad de *futuras* estructuras conceptuales para capturar el hecho cualitativo. Basta un simple cambio (ampliación, desviación semántica, etc.) en la referencia de ciertos conceptos científicos –lo que ocurre continuamente no sólo en ciencia, sino en el lenguaje ordinario– para que este dictamen deje de parecer plausible.

Nos queda b2): La ciencia deberá generar nuevas estructuras conceptuales, o nuevos modos de acceso o incluso una nueva noción de “empírico” para poder capturar la

natural»: la presuposición de una conexión entre el noéma y la existencia de un objeto trascendente a ese noéma. Pero esa actitud natural es posterior a la actitud propia de la fenomenología trascendental, que consiste en una epojé acerca de cualquier rasgo que implique trascendencia. En otras palabras: la brecha solo puede aparecer en una actitud que es *posterior* al propio noéma; i.e., en una actitud hacia los noémas en la que se ponen objetos trascendentes que no están presentes en el propio noéma. Esta “posición” de objetos trascendentes es parte de un cierto carácter tético de la experiencia. La epojé debe neutralizar ese carácter tético. (Cf. San Martín Sala, 1987, págs. 63 – 68). No discutimos esta solución en el cuerpo del texto, porque nos interesa remarcar la relación entre ciertos planteos de la modernidad y planteos aparecidos en los siglos XX y XXI, aunque la propuesta que realizamos al final de este trabajo tiene algunos puntos de coincidencia con las respuestas de la fenomenología clásica.

especificidad del quale. Parece legítimo exigirle a la ciencia el desarrollo de nuevas modalidades de investigación, resemantizaciones conceptuales y demarcaciones. El problema es determinar el límite entre una demanda científica genuinamente plausible y una demanda desmedida e incluso incompatible con la investigación científica natural actual. Quien dictamine que *todo* el arsenal conceptual de la ciencia natural es *radicalmente inadecuado* para capturar la esencia de los fenómenos cualitativos conscientes e incorporarlos en la ciencia natural, debería poder indicar en algún sentido qué clase de estructura conceptual sería aceptable adoptar y cómo podría llevarse a cabo un programa de tales características. De otro modo, es difícil distinguir la insatisfacción psicológica, puramente subjetiva⁴³, de una auténtica crítica a los métodos explicativos actuales. Es necesario analizar hasta qué punto es exigible una explicación de los estados de conciencia y *qué* de ellos debe ser explicado para que cuente como explicación *suficiente* del fenómeno.

Una de las hipótesis que sirven de hilo conductor entre esta sección y la sección anterior es que *muchos de los intentos por explicar los fenómenos cualitativos de la conciencia apelan a nociones que contienen resabios de vaguedad o ambigüedad propias de las conceptualizaciones modernas del realismo por representación*. Así como notamos disunidad y polisemia con el término *idea* de la modernidad, también existe una disunidad con respecto a *lo que es ser una característica fenoménica de la conciencia*: se han tomado muchas veces como sinónimos los términos *parecer/aparecer/ apariencia* y *qualia*, cuando en realidad es deseable hacer importantes distinciones entre ellos, algunas de las cuales son paralelas a la distinción que hacíamos dentro de la noción de *idea* moderna en su aspecto puramente *fenoménico* y en su aspecto de *fenómeno que indica su propia causa externa*. Así como notamos que, en el intento por reducir cualidades secundarias a cualidades primarias se introducía un elemento fenoménico en el propio concepto de cualidad primaria que hacía necesario distinguir cualidades primarias *reales* de cualidades primarias *a secas*,

⁴³ Sugerida por el § 6.53 del *Tractatus* wittgensteiniano: “El método correcto de la filosofía sería propiamente este: no decir nada más que lo que se puede decir, o sea, proposiciones de la ciencia natural –o sea, algo que nada tiene que ver con la filosofía-, y entonces, cuantas veces alguien quisiera decir algo metafísico, probarle que en sus proposiciones no había dado significado a ciertos signos. Este método le resultaría insatisfactorio – no tendría el sentimiento de que le enseñábamos filosofía-, pero sería el único estrictamente correcto” (1999, pág. 183)

también aparecen importantes (y cruciales) elementos fenoménicos en la explicación funcionalista reductiva de la actualidad. Otro de los puntos importantes (y parcialmente derivado del anterior) es que *durante los siglos XX y el XXI no se ha podido tender el puente entre la conciencia y la realidad externa*. La reducción de los fenómenos de conciencia a sus causas no fenoménicas no ha podido realizarse. En esta segunda sección analizaremos algunas de las investigaciones que proponen una explicación funcionalista de los términos fenoménicos.

Los juicios fenoménicos son aquellos en los cuales el sujeto explicita, mediante una proposición, cuáles son o han sido sus estados fenoménicos. Es común que estos juicios sean introducidos mediante el verbo “parecer”: “Me parece que está haciendo frío”; “Parece que esta mesa es de color verde”. El juicio fenoménico es aquel en el que se explicita lo que está ocurriendo en la conciencia del propio individuo que está realizando el juicio. En § 2.1, 2.2, 2.3 y 2.4 analizaremos el componente fenoménico de este juicio, y evaluaremos este componente a la luz de la posición reductivista de Daniel Dennett, quien intenta zanjar la cuestión de la reducción entre fenómenos y acaecimientos mediante la estrategia de *negar* la existencia de los fenómenos de conciencia. Mostraremos que este método reductivista - eliminativista comete una petición de principio y que, por lo tanto, no resulta viable. En § 2.4 expondremos cuáles son los requisitos que debe satisfacer la explicación de un hecho para que dicha explicación cuente como relevante y suficiente, y para que no deje resto explicativo. Los reductivistas y los no reductivistas tienen un desacuerdo con respecto a qué cuenta como *explicaciones conceptualmente adecuadas* (CAE) del ámbito fenoménico. De acuerdo a los tipos de requisitos que se imponen para satisfacer estas CAEs, las estrategias explicativas serán divergentes. En § 2.5 estableceremos las diferencias entre el juicio fenoménico (el juicio de parecer) y las características normalmente atribuidas a los fenómenos de conciencia (los *qualia*). Dado que para los programas reductivistas las CAE cumplen su cometido cuando reducen el ámbito fenoménico al ámbito de los acaecimientos, examinaremos las posiciones de Jeff Foss (§ 2.6), Adam Pautz y Alex Byrne & Michael Tye (§2.7), y Mohan Matthen (§ 2.8) con respecto a la reducción de los fenómenos del color a sus elementos funcionales. Mostraremos que, en todos los casos en los que se da cuenta del éxito de un programa reductivista, en

realidad tal reducción ni siquiera se ha operado y por lo tanto la explicación del hecho no está dada. Concluiremos que los programas reductivistas, tanto los actuales como los de la modernidad, al intentar reducir el aspecto fenoménico al aspecto funcional o corpuscular (o a lo que fuera que cuente como acaecimiento externo de base), no pueden evitar la reaparición del aspecto fenoménico supuestamente reducido en su *explanans*. El reductivismo obtendría su éxito si el aspecto reducido (el fenoménico) desapareciera completamente en la explicación conceptualmente adecuada del hecho. Pero si este aspecto vuelve a aparecer una y otra vez en el *explanans*, entonces la reducción no se ha operado. En § 2.9 extraeremos las consecuencias de de los apartados anteriores. La consecuencia más notoria es que las definiciones de “visión de color” que dejan afuera el aspecto fenoménico, no pueden establecer una explicación adecuada y suficiente de la noción de color en tanto que percibido: de algún modo, parece que la apelación al fenómeno es necesaria para poder realizar la propia explicación reductiva (lo cual, paradójicamente, da por tierra con el éxito de la reducción) . Otra consecuencia importante es que la propia noción de “color”, sin referencia al fenómeno cualitativo, parece referirse a un sinnúmero de acaecimientos del mundo que no son ni homogéneos, ni fácilmente definibles. Entre esta noción de “color” sumamente disunificada y la noción de fenómeno de color, resulta implausible que se pueda establecer algún tipo de correlación directa. Mucho menos plausible parece que se llegue a lograr una reducción entre ambos ámbitos.

Algunas consideraciones terminológicas: cuando hablamos de características fenoménicas, nos podremos referir a características de los siguientes tipos:

- a) Características globales o “cómo es ser”: siguiendo una modalidad popularizada por Nagel (1974), un estado consciente de un determinado individuo se caracteriza por *ser como* ese individuo en ese estado. Esta caracterización es *global* en el sentido de que trata de capturar la esencia de *todo* el campo consciente de un individuo en un momento determinado, y no de un hecho consciente puntual. Si se pregunta por la característica global de una conciencia se está inquiriendo por la específica subjetividad del ser-consciente-en-tanto-individuo-de-tal-especie-y-tal-circunstancia: ¿Cómo es ser murciélago? ¿Qué se siente ser un animal de otra especie? En versiones especulativas más radicales,

las características globales pueden preguntar por el ser-como ya no de individuos conscientes sino incluso de seres que son de hecho no conscientes. Una típica pregunta global de este tipo es: ¿cómo es ser un termostato? (Chalmers, 1999, pág. 371)

- b) Características locales o puntuales: Mientras en la caracterización global se pretende una explicación de la totalidad del campo consciente, en lo que respecta a la caracterización local sólo se pregunta por un determinado aspecto de este campo consciente. Desde este punto de vista, “¿cómo es ver rojo para la especie X?” es una pregunta por la fenomenología local y no por todo el campo consciente de los individuos-X.

En lo que sigue nuestro análisis será sobre una característica local: la visión del color, y ocasionalmente hablaremos de características globales.

2.1. Acerca del parecer

If you believe something obvious in the absence of much in the way of reasons, you might be epistemically in the clear. But if you believe something that is not obvious in the absence of reasons, that is not so good – John Gibbons.

Los modernos llamaban ‘idea’ contenido fenomenológico de la conciencia. Vimos que, en el programa mecanicista de Locke, las ideas no podían explicarse aplicando una reducción de los componentes fenoménicos a las cualidades primarias reales de la materia, puesto que estas cualidades no eran asequibles a la ciencia humana. A su vez, las ideas provenientes de cualidades secundarias parecían tener un componente fenoménico irreductible, anómalo, cuya explicación podía residir en la arbitraria decisión divina de unir un determinado acaecimiento a un determinado tipo de idea. La brecha entre ideas y mundo parece infranqueable. Sin embargo, la modernidad cartesiana mantiene que el sujeto tiene autoridad acerca de sus propias ideas, y en ese sentido la propia fenomenología es un ámbito seguro. Puedo emitir el juicio ‘me parece que la pared es blanca’ con la certeza que me otorga el ser dueño de mi propia fenomenología. Pero, ¿qué ocurriría si en verdad nuestra conciencia no es dueña de sus ideas o si, mejor dicho, no estuviera en posición privilegiada con respecto a los juicios fenoménicos cuyo contenido es su propia fenomenología? En ese caso, el sujeto no sólo no puede conocer con certeza ‘lo que hay afuera’, sino ni siquiera ‘lo que hay adentro’ de su propia conciencia. En la primera parte analizamos que, dentro del marco cartesiano, las proposiciones cuyo contenido es un acaecimiento (“Hay una mujer cruzando la calle desierta”) resultaban problemáticas, porque no podía garantizarse la conexión entre el contenido de conciencia (la idea de una mujer cruzando la calle en mi mente) y el hecho externo. De modo que sólo podíamos realizar, legítimamente, proposiciones cautelosas del tipo “me parece que hay una mujer cruzando la calle”, en las que lo único que se afirma es el hecho mental (el “parecer”). El verbo ‘parecer’ nos indica que estamos haciendo un juicio fenoménico. En este capítulo analizaremos la validez de este juicio. Querremos mantener que, a pesar de las críticas, el juicio fenoménico puede mantenerse en pie, a pesar de diversos intentos por eliminarlo o reducirlo.

El problema de la autoridad en primera persona con respecto a los juicios fenoménicos puede ilustrarse con el diálogo de Raymond Smullyan titulado “Pesadilla epistemológica” (Smullyan, 1983)

En la pesadilla epistemológica, Frank afirma ante un epistemólogo que *le parece* ver rojo:

Epistemólogo: ¡Mal!

Frank: No creo que hayas oído bien lo que dije. Dije simplemente que a mí me *parecía* rojo.

Epistemólogo: te oí y te equivocaste.

Frank: Quiero aclarar bien esto. ¿Quisiste decir que me equivoqué al decir que el libro es rojo, o bien que me equivoqué porque me *parece* rojo?

Epistemólogo: Obviamente no pude haber querido decir que te equivocaste en el sentido de que *es* rojo, puesto que no dijiste que es rojo. Todo lo que dijiste es que te *parece* rojo a ti, y es *esta* la afirmación que está equivocada.

Frank: Pero no puedes decir que la afirmación “a mí me *parece* que es rojo” sea equivocada. (...) ¡Sin duda *yo* sé de qué color me *parece* el libro!

Epistemólogo: vuelves a equivocarte (Dennett, Daniel; Hofstadter, Douglas, 1983, pág. 538)

La discusión del diálogo es acerca de la legitimidad del uso del verbo *parecer* dirigido hacia la primera persona. ¿En qué momentos utilizamos de ese modo el verbo *parecer*?

En muchos casos, lo hacemos para emitir una opinión, marcando el hecho de que se trata un punto de vista no suficientemente justificado: “Me parece que hoy va a llover”; “Parece que Rebeca ama a José” o “Me parece que estás actuando mal”. En este sentido, puede entenderse como sinónimo de “creer”. Por lo general, este uso de “parece” se justifica por algún indicio, aunque dicho indicio no es concluyente.

A veces lo utilizamos para referirnos a nuestras sensaciones o nuestras impresiones acerca de ellas: “La salsa me parece sabrosa”; “El anillo parece rosa”, “El círculo parece verde y parece estar girando.

De los dos usos de “parecer” nos interesa especialmente el segundo. ¿En qué circunstancias una persona dice “Me parece rojo” y no, directamente, “Es rojo” o “Lo veo rojo”? Aunque las circunstancias son variables, hay sin duda un criterio general de uso. A veces, cuando por un efecto de iluminación no estamos seguros acerca de si la

pared es blanca o celeste, decimos “me parece celeste (o blanca)”, remarcando que no estamos en una condición epistémica óptima para decidirlo: nuestro “parecer” es una inferencia, como cuando vemos un cartel a lo lejos y no podemos distinguir con claridad lo que dice. En esos casos, arriesgamos (mediante una inferencia) una interpretación de lo que *no* podemos leer. Este caso puede ilustrarse a partir de un ejemplo de Dretske citado por Chalmers (1999, pág. 297): “Usted me muestra siete dedos y yo veo los siete. Pero no tengo tiempo para contarlos y erróneamente supongo que veo ocho dedos. De modo que juzgo que hay ocho dedos ante mí, pero mi experiencia fenoménica es de siete dedos” (Dretske, 1995). Queda claro que el juicio “me parece que hay ocho dedos” se basa en una experiencia defectuosa, pero no informa de esa experiencia sino de una inferencia a partir de dicha experiencia defectuosa. Se trata de una suposición y no del *informe directo* de dicha experiencia.

Otras veces, en cambio, utilizamos el “me parece” como un sinónimo de “percibo” o “tengo tal o cual estado de conciencia”: “Me parece que hay un círculo verde”; “Parece que comienza a doler nuevamente”. En este caso, el verbo no marca la inseguridad del hablante con respecto al hecho aparente; lo que marca es que quizás, a pesar de estar percibiendo esa sensación, es posible que no haya algo que la provoque, o que la causa de la experiencia tal vez no sea la que se supone inicialmente. No es una duda con respecto a la sensación en sí, sino con respecto a su causa: “Parece que veo un anillo rosa, pero quizás se trata de una ilusión óptica”. ¿Por qué alguien podría inferir que eso que está viendo no existe o es diferente a como lo ve? Puede haber miles de razones; la más usual es que, cuando alguien nos da un argumento para descartar nuestra experiencia, utilizamos el cauteloso “parece” para referirnos a ella: “Te vi ayer por la calle. Llevabas un sombrero verde”, decimos. “Qué curioso. Ayer yo no llevaba sombrero.”, replica nuestro interlocutor. “Bueno, me pareció que lo llevabas”. En este caso es posible que hayamos añadido (falsamente) la representación del sombrero verde a la percepción de esa misma persona que nos cruzamos, o bien que hayamos confundido a nuestro interlocutor con otra persona que *de hecho* sí tenía un sombrero verde. En cualquiera de los dos casos, no estamos poniendo en duda nuestra experiencia de “ver un sombrero verde”, sino la atribución de que ese sombrero la tenía nuestro interlocutor. Aun cuando en ningún momento haya habido un sombrero

verde (quizás fue añadido por la imaginación), el “me pareció” utilizado enfatiza el hecho de que X lo llevara, no la sensación compleja del sombrero verde (aunque, en algunos casos, podemos preguntarnos adicionalmente: “¿de dónde habré sacado la idea de que X tenía un sombrero verde?”)

Otra razón poderosa para dudar de la existencia de lo que se percibe es que, en muchos casos, somos víctimas de ilusiones sensoriales y perceptivas. En estos casos, el “me parece” se utiliza para añadir cautela a una sensación que *de hecho* se está notando, pero no expresa duda acerca de dicha sensación. Analicemos la relación entre el parecer y la sensación en la figura siguiente: (Figura 6)

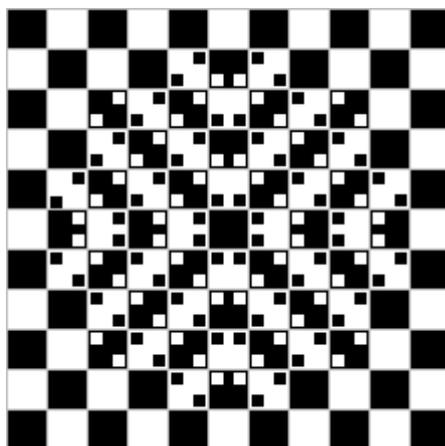


Figura 6

En una observación descuidada (¡E incluso en una un poco más cautelosa!) creeríamos que las líneas ortogonales que forman las cuadrículas están ligeramente curvadas. Si medimos con una regla dichas líneas, nos daremos cuenta de que *no* hay curvas. Ahora bien, si se nos pregunta qué estamos viendo en este dibujo, diremos que se trata de una cuadrícula, con figuras rectangulares blancas y negras en dos tamaños, distribuidos en forma simétrica, cuyas líneas tienen una notable curvatura⁴⁴. Nuestra percepción inmediata nos suministra de modo patente, persistente e inequívoco, la

⁴⁴ Este informe, desde luego, podría diferir con respecto al propuesto. Un informe alternativo llamaría la atención acerca de una cierta incongruencia entre la “apariencia general” (curvada) y la apariencia de cada cuadrado en particular. Mientras que cada cuadrado visto de modo individual parece, efectivamente, cuadrado, la apariencia en un golpe de vista general hace “aparecer” las curvas. Dejamos de lado este informe posible, porque en un caso así el informante tendría motivos de sospecha a partir de esta incongruencia. Para nuestro argumento, necesitamos que esa incongruencia sólo se revele en un estado posterior al juicio inmediato y no en el mismo momento en que se emite el juicio.

“sensación de curva”. Quienes están acostumbrados a las ilusiones ópticas sospechan desde el primer instante que hay algo engañoso, y en esos casos sus informes revelan cierta cautela desde el primer juicio: “Me parece que las líneas de la cuadrícula están curvas”, o “parece *como si* las líneas estuvieran curvas”. Pero si no hay razones para sospechar, se dirá que las líneas *son* curvas. En algunos casos, aun si no hubiera motivos para la sospecha, se podrá decir que “parecen curvas”, pero ese “parece” no pretende ser un juicio de segundo orden con respecto al informe de la sensación; simplemente se lo utiliza como sinónimo de “tengo la sensación X”. Retomaremos esta ambigüedad del “parece” en el parágrafo 2.2.

Si comparamos los dos estados: el momento en el que se afirma, sin lugar a dudas, que “las líneas son curvas” (estado A) con el estado posterior en el que se llama la atención acerca del hecho de que las líneas, en realidad, *no son* curvas (estado B), podremos decir algo como esto: “Me había parecido que las líneas eran curvas. Tenía toda la certeza de ello. Ahora veo que mi parecer era equivocado: en realidad *no son* curvas”. Pero esta ilusión óptica, como muchas otras, tiene un curioso efecto: *sigue pareciendo lo que no es, incluso cuando sabemos cómo es realmente*.

Chisholm (1982, pág. 39) distingue entre “aparecer” (*appearing*) y “parecer” (*seeming*). El primero para hablar acerca de algo inmediatamente percibido (“Las líneas se aparecen curvas”); el segundo para referirse al parecer puramente inferencial (“Aunque me parece que las líneas no son curvas”).

Si en el estado A hemos afirmado sin dudar: “Las líneas son curvas”, lo que hemos dicho es un informe directo de nuestro parecer inmediato: informamos *lo que se nos aparece*. Alguien podría aclararnos que no hubo nada extraño en la percepción; que los órganos perceptivos en realidad hicieron bien su trabajo (percibieron “lo que debía ser percibido” y no lo que yo dije que vi). En ese caso, podríamos aclarar: “Cuando en el estado A dije ‘las líneas son curvas’, simplemente dije *lo que me había parecido a mí*.” Con este expediente, lo que hemos hecho fue equiparar el *parecer* con nuestra *convicción espontánea* y persistente de que tuvimos una determinada sensación. No estamos diciendo que estábamos equivocados con respecto a la sensación; simplemente habíamos hecho un juicio incorrecto, al decir, sin ningún matiz, que “las

líneas son curvas”, cuando en realidad deberíamos haber dicho que “se nos aparecían” curvas. Este es el punto que nos interesa: el *parecer*, cuando se refiere a una sensación notada (a un *aparecer*), se presenta como esencialmente *incorregible*. Las ulteriores correcciones sólo afectan al juicio, no a “lo que nos había aparecido” inmediatamente.

En el estado B hemos ganado algo: sabemos que lo que se nos *aparece* no *es*. Es un punto importante si lo que nos interesa es establecer las relaciones entre nuestras sensaciones y los hechos “reales”, esto es, los hechos tal como los describe una teoría científica vigente. Cuando caemos en la cuenta de que eso que percibimos *no es*, estamos en condiciones de enfocar la atención de una manera diferente. Veremos que los cuadrados más pequeños están dispuestos de una manera tal que, quizás, condicionan *el modo en que se nos aparece* el contenido global de la figura. Miraremos cada una de las figuras blancas y negras, y constataremos que, tomadas de modo individual, cada una es efectivamente cuadrada. Ahora, una parte de nuestro parecer está “infectado” por el dato (corroborado a partir de las figuras individuales, o mediante una regla superpuesta en las líneas del dibujo) de que *no hay* líneas curvas. Otra parte, sin embargo, seguirá bajo el hechizo ineliminable de la ilusión óptica.

2.2. ¿No le parece que carece de parecer?

Cuando vemos una superficie roja o sentimos el olor de una rosa simplemente informamos acerca del hecho percibido sin referencia a nuestro parecer. Ni siquiera hacemos mención acerca del juicio perceptivo en sí mismo (no decimos “estoy percibiendo una superficie roja” o “se me aparece roja”, sino “la superficie es roja” o “hay una superficie roja”). Estos juicios, en los que simplemente informamos del contenido de nuestra conciencia sin referirnos al acto perceptivo en sí mismo, son juicios de primer orden (Chalmers, 1999, pág. 281). Esta clase de juicios posee una fenomenología que está inmediatamente proyectada en el hecho externo. Sólo en un segundo nivel (un juicio de segundo orden) examinamos ese juicio (por ejemplo, si alguien nos pide que demos cuenta de ese color que hemos visto porque puede haber

dudas al respecto) y diremos “me pareció que la superficie era roja”. Supongamos que alguien nos insiste para que aseguremos si efectivamente la superficie era roja o naranja. Quizás había poca luz, lo que condicionaría nuestro juicio acerca del color. En ese caso, simplemente diremos “dadas las condiciones de luz (y otras que se evaluarán), me pareció que la luz era roja.”

En este juicio de parecer no estamos dando simplemente un juicio fenoménico; es un juicio acerca de nuestro juicio acerca de un fenómeno. El fenómeno original era, simplemente, la rojez de la superficie *tal como nos había aparecido inmediatamente*. Este segundo juicio, en cambio, es un juicio correctivo acerca *de nuestra visión* de la rojez y es por ello un juicio de segundo orden. La utilización del verbo “parecer”, en este caso, incorpora una cautela o una neutralidad con respecto a la existencia efectiva del contenido del parecer inicial. El parecer en su sentido fenoménico sólo se da en el juicio de primer orden, *aun cuando no lo expresemos como un parecer*. Decimos “La superficie es roja” y con ello *expresamos* (aunque no decimos) nuestro parecer. Dicho en términos de Wittgenstein, el parecer se *muestra* cuando *decimos* lo que percibimos. Llamemos a este parecer no dicho, implícito en los juicios de primer orden, “parecer₁”, para distinguirlo del parecer de segundo orden, “parecer₂”

¿Es posible que una persona esté desacoplada de su propio parecer₁? Supongamos que alguien está viendo una superficie roja, pero no le parece₁ que ve una superficie roja. Si en el acto de percibir ya hay un parecer₁ implícito (un parecer que se muestra aunque no se diga), la situación planteada no es posible. Que yo perciba algo en este momento equivale a que me parezca₁ que lo estoy percibiendo. Ese “parecer₁” es un acto de percatación consciente inmediata, una convicción espontánea, y resultaría curioso que alguien dijera: “Estoy viendo rojo, aunque no *me parece₁* verlo”. Si se dijera algo así, se podría preguntar por las características de ese “ver rojo”. Si quien emite tal enunciado no puede dar características fenomenológicas de la superficie roja que está viendo, podría concluirse que en un sentido importante no está viendo. En ese caso, es correcto que no le parezca₁ que está viendo rojo, aunque sería falso que está viendo rojo. Por el contrario, si puede dar características fenomenológicas, debemos concluir que aunque diga que no le parece₁, en realidad le está pareciendo₁. Pero este caso es desconcertante, porque el parecer₁ es idéntico a la convicción

espontánea e inmediata de una sensación: resultaría por lo menos incongruente que una persona enuncie el contenido de una sensación que está teniendo pero no tenga la convicción espontánea de que está teniendo esa sensación, como si el fenómeno de la sensación le fuera transmitido por un medio indirecto. Por eso, no parece plausible que a una persona (a menos que esté desconectada de su propia fenomenología) no le parezca que está percibiendo algo que de hecho informa como si lo estuviera percibiendo.

El fenómeno de la visión ciega (Humphrey, 1995; Dennett, 1995) podría suministrar alguna afirmación desconcertante como la que supusimos arriba. La visión ciega consiste en un defecto neurológico que afecta a la fenomenología de la visión. El sujeto que la padece afirma que no puede ver, aunque en algunos casos se comporta como si de hecho estuviese viendo. Si a un afectado de visión ciega se le pide que seleccione una ficha de determinado color, la selección será correcta –al menos en una cantidad de veces significativamente superior a la esperable en una elección azarosa– a pesar de la insistencia del sujeto en que no puede ver. Una persona a la que se le da a entender que selecciona los colores correctos a pesar de que dice que no puede verlos, podría afirmar con certeza que en algún sentido (no fenomenológico) ve el color rojo, aunque no le parece₁ que lo vea. El parecer₁, tal como lo estamos analizando, está fuertemente atado al contenido fenomenológico. Es extensa la literatura acerca de experimentos mentales sobre la posibilidad lógica de la existencia de zombis, robots u “hombres de imitación”: seres que no poseen estados fenomenológicos, aunque funcionalmente se comportan como si los poseyeran (Chalmers, 1999, pág. 132): los robots (que serían duplicados funcionales idénticos a las personas de carne y hueso) pueden decir que les parece que ven rojo, aunque en realidad no tienen el estado cualitativo de conciencia propio de una auténtica persona cuando ve rojo. A pesar de la ausencia de estados fenoménicos, ¿tiene el robot algún tipo de parecer₁? ¿Cuál es la relación que tiene un robot con sus pareceres? Si el parecer₁ está fuertemente atado a un estado fenoménico, entonces el robot no tendrá parecer₁ y sí parecer₂. En el parecer₂, recordemos, estamos haciendo un juicio de segundo orden acerca del fenómeno, y en este caso el contenido es el juicio de primer orden (en el cual está incluido el fenómeno). El robot puede efectuar juicios de

segundo orden (parecer₂) pero no puede hacer juicios de primer orden (parecer₁). De modo que cuando dice “me parece que esta mesa es roja”, el robot cree (o tiene la intención) de referirse *directamente* al fenómeno de la rojez de la mesa. Pero él no tiene acceso a ese fenómeno, así que su parecer está mal dirigido: en realidad él está efectuando un juicio de segundo orden (parecer₂) en el cual el fenómeno no está presente de modo inmediato. Cuando el robot dice “me parece que la mesa es roja” no está teniendo la convicción espontánea de que la mesa es roja, sino informando acerca de una inferencia, o simplemente está adivinando. ¡Lo curioso es que a él no le parece que está adivinando! Si pudiera informar acerca de su parecer, dirá que *le pareció que estaba viendo rojo* (queriendo decir con esto que, cuando percibió la mesa, tuvo la convicción inmediata y espontánea de que era roja), pero en realidad estará diciendo que le pareció₂ que le parecía₁ que veía rojo. Dado que no puede tener una experiencia fenoménica, en rigor no pudo parecerle₁ que veía rojo: todo lo que tiene es un juicio de segundo orden acerca de una convicción inmediata y espontánea que estaba ausente. Si le “parece₂ que le parece₁”, en realidad está teniendo una *falsa apariencia de fenomenología*, porque todo lo que hay en su estructura funcional son pareceres₂. Él puede decir que le parece, aunque no le parezca, y que todo en su estructura funcional se comporta como si le pareciera. ¡Todo en él se comporta como si le pareciera que ve rojo, incluso es capaz de comparar muestras de colores, pero en realidad no puede parecerle que ve rojo porque no ve rojo! El robot está preso de la pesadilla epistemológica de Smullyan: no es dueño de sus propios juicios de parecer. Quizás una máquina, o nosotros mismos (quienes nos jactamos de no ser robots) podamos informarle acerca de la corrección de sus juicios de parecer. Por eso el robot, o Frank pueden decir:

Frank: ¡Sin duda yo sé de qué color me *parece* el libro!

Epistemólogo: Vuelves a equivocarte.

Frank: Pero, nadie puede decir cómo me parecen *a mí* las cosas.

Epistemólogo: Lo siento, pero otra vez estás equivocado.

Frank: Pero, ¿quién puede saberlo mejor que yo?

Epistemólogo: Yo. (...) Por la lectura directa de los pensamientos de una persona.

Frank: ¿Quieres decir que tienes poderes de telepatía?

Epistemólogo: Claro que no. Hice sencillamente lo que era obvio hacer. He construido una máquina de leer el cerebro –conocida bajo el nombre técnico de cerebroscopio– que está funcionando en este momento mismo en este cuarto y escudriñando cada

célula nerviosa de tu cerebro. Así puedo leer todas tus sensaciones y pensamientos, y es una verdad objetiva que este libro *no* te parece rojo.

Frank (muy impresionado): ¡Vaya, habría jurado que ese libro me parecía rojo! ¡Te juro que parece parecerme rojo!

Epistemólogo: Perdona, pero volviste a equivocarte. (1983, págs. 538-9)

Frank está en la preocupante posición de no ser dueño de su parecer. El robot no es dueño de sus pareceres en el sentido de que no tiene auténticos pareceres, aunque en sus juicios se esfuerce por convencernos de que los tiene (y aunque diga que le preocupa, en verdad no le preocupa). ¿Puede el parecer ser objeto de un análisis en tercer persona, tal como lo realiza el epistemólogo? ¿Es posible extraer una “verdad objetiva” acerca del parecer?

Sabemos que el contenido del parecer del robot no es “real” y por lo tanto no tiene una “apariencia real”, sino la apariencia de una apariencia. No se le *aparece* algo como rojo, aunque diga que sí se le aparece. Más abajo, en § 2.3 analizaremos la noción de “apariencia real”. Por ahora, tendremos en cuenta que un juicio de parecer está conectado con un fenómeno y con un sujeto consciente para el cual ese fenómeno está dado. Quien afirmara con toda sinceridad sus pareceres sin que haya un estado fenoménico concomitante, estaría sistemáticamente desconectado de sus propias percataciones. En principio, el Frank de la pesadilla epistemológica *no es un robot* y sus percepciones son normales, de modo que es posible –si se acepta la idea de que es posible escanear el contenido de una mente para decidir lo que le parece- estar equivocado con respecto al parecer aun cuando a uno no le parezca que esté equivocado. La máquina del epistemólogo ofrece un método indirecto para conocer el parecer de una persona. Quizás aquí está la pieza del engranaje retórico de esta pesadilla, y gran parte de su fuerza: proponer que una máquina puede decidir acerca de *nuestro* propio parecer implica que *nuestra* convicción inmediata y directa de una sensación no es esencialmente nuestra, sino que está sujeta a un análisis de tercera persona. Humphrey (1995, pág. 149) afirma que las sensaciones “hablan por sí mismas o delatan sus propiedades características, de modo que el sujeto se percata de ellas en forma directa e inmediata” ¿Llamaríamos *nuestro parecer* al parecer *sometido a una máquina que nos lo informe*? Definitivamente, ese no es *nuestro* parecer. El parecer₁

está indexicalmente atado a un estado fenoménico y al *acceso inmediato y directo* de una determinada conciencia a ese estado. Si a ese parecer se lo somete a un método de evaluación indirecto, tal parecer ya deja de ser nuestro. ¿Qué cambiaría si se nos dijera que, en realidad, el libro nos parece verde en lugar de rojo? Si mi convicción inmediata me indica que es rojo, de nada vale que desde el acceso de tercera persona se me diga que me parece verde (en ambos casos, tanto en mi parecer como en el diagnóstico de la máquina, esto es independiente del hecho de *de qué color es efectivamente el libro*)

Imaginemos un caso paralelo. Se construye una máquina que permite decidir si determinados actos pertenecen o no a nuestra voluntad. Nos sometemos voluntariamente a la máquina, y la máquina dictamina que no lo hemos hecho voluntariamente, o que no éramos del todo conscientes cuando decidimos someternos a ella (en rigor, no lo *decidimos* nosotros)⁴⁵. Como en el caso de la pesadilla epistemológica, la máquina dictamina, ante cada una de nuestras decisiones, que no las hemos llevado a cabo voluntariamente. Desde la acción más trivial hasta la más concienzuda; desde mover un dedo al compás de la música hasta hacer una larga carrera universitaria. Aun cuando reconstruyéramos la extenuante cadena de

⁴⁵ Aunque no tiene que ver con nuestro propósito, es interesante especular sobre las consecuencias del caso contrario al que supusimos. Supongamos ahora que quien se somete a la máquina es una persona que actúa pero no “siente” que lo hace voluntariamente. Por ejemplo, cuando levanta un brazo o cuando camina, no “siente” que él decide levantar el brazo o caminar, como si su cuerpo se moviera “bajo las órdenes” de otro. Lo que esta persona querría saber es si su “sensación de no hacer las cosas voluntariamente” es legítima. Supongamos que la máquina dictamina que, a pesar de que no “siente” que hace las cosas voluntariamente, en realidad sí las está haciendo. Eso significa que hace las cosas “bajo sus órdenes”, y no “bajo las órdenes de otro, o de nadie”. En este punto se podría preguntar: ¿se mueve “bajo sus propias órdenes”? ¿Cómo son esas órdenes? ¿Cómo es sentir a tu cuerpo moverse bajo tu voluntad? ¡Supongamos, también, que tu cuerpo se mueve tal como querías que se moviera, pero no tienes la sensación de que se está moviendo por tu voluntad! Quieres levantar el brazo y el brazo se levanta, pero *como si no fueras tú el que lo levanta*. (Cf. Wittgenstein, [1988] § 417- 8). En la historieta “El Eternauta” ocurre algo parecido a lo especulado aquí: Juan Salvo es capturado por extraterrestres y – hasta donde él supone- es convertido en robot. A pesar de que tiene pensamientos conscientes, *sabe* (o cree saber) que son los pensamientos de un robot. Intenta mover la mano y la mano se mueve, pero de inmediato supone que sólo se ha movido por las órdenes del *Mano*, el extraterrestre que lo manipula. A medida que su cuerpo parece responder a sus órdenes (y no a las del *Mano*), Juan Salvo se da cuenta de que es dueño de su voluntad y de su cuerpo. Pero quizás el Mano le hizo creer que sus movimientos eran voluntarios. Quizás parte de ser un robot consista en creer que no se es robot. La extrañeza inicial de Salvo para con su propio movimiento de manos ilustra esta posibilidad de “levantar el brazo *como si no fueras tú el que lo levanta*”. Al respecto, hay estados alterados interesantes, por ejemplo en la esquizofrenia o en estados histéricos profundos: el sujeto realiza actos voluntarios de mover partes del cuerpo, pero informa tener la experiencia de no haber sido él quien realmente las movió.

razonamientos que (supuestamente) nos llevaron a tomar tales decisiones, la máquina dictaminaría que todos esos razonamientos -al igual que las acciones que supuestamente ejecutamos a partir de ellos- no eran voluntarios. Yo tuve la sensación de que decidía voluntaria y conscientemente, pero en realidad era títere de decisiones que no emanaban de mi voluntad ni de mi conciencia. Al igual (o quizás peor) que el Frank de la pesadilla epistemológica, no soy dueño de conocer qué cosas decido yo y qué cosas se deciden solas bajo la apariencia de que son decididas por mi voluntad.

En ambos casos, tanto el del parecer como el de la voluntad, existe un requisito indexical mínimo necesario para la correcta utilización de los términos. Cuando digo que algo me parece rojo o que hice una acción de manera voluntaria, no necesito recurrir a fuentes indirectas para corroborar la verdad de lo que digo (Dejemos de lado la pregunta acerca de qué es lo que debería computar una máquina para decidir acerca de nuestra voluntad o acerca de nuestro *auténtico* parecer). Si necesitara de tales fuentes, no puedo esperar que esas fuentes me digan algo acerca de mis juicios que no podría decir yo mismo. En todo caso, si una fuente tal me dice algo que es incompatible con la relación que tengo con mis estados conscientes (mi parecer, mis decisiones, mis creencias), eso pondría en duda ya no mi parecer o mi voluntad, sino una cualidad fundamental de los propios estados conscientes: su accesibilidad inmediata, su autoevidencia. No es necesario que haga una inferencia para saber que estoy teniendo una determinada sensación: la propia sensación es el método para saber lo que me está pasando. Un método indirecto para conocer el estado de mi propia fenomenología es precisamente la negación de la fenomenología. Si hay que preguntar, algo anda mal. Hilary Putnam ([1960] 1984, pág. 113) establece que “La pregunta «¿cómo sé que me duele» es una pregunta desviada («lógicamente rara»”, en el sentido de que se desvía de la regularidad semántica usual del término “dolor” (No hay contextos de uso claros en los cuales una persona le pregunte a otra si realmente le duele una parte de su cuerpo o no). La única respuesta esperable ante una pregunta de ese tipo es: “doliéndome”. Si juzgo que me parece₁ que me duele, entonces me duele, a menos que no exista una fenomenología y mi juicio de parecer₁

no tenga un contenido cualitativo⁴⁶. Es posible que el contenido cualitativo no sea algo evidente, y que algunos se propongan negarlo, al menos en su sentido fenomenológico usual. Examinaremos a continuación la posición *negacionista* de Daniel Dennett acerca de este problema. Porque quizás todo el enredo que surge a partir de la irreducibilidad de los estados fenoménicos se deba a que, de hecho, tales estados no existen.

2.3 Hablando acerca del inexistente anillo rosado del que no tiene sentido hablar.

En los dos apartados anteriores se examinó la relación entre nuestro parecer y la autoridad de primera persona para juzgar acerca de ellos. En los juicios de parecer se informan las características fenomenológicas de un determinado hecho de conciencia. Si a un sujeto se le quita la autoridad para hablar acerca de su propio parecer, entonces su conciencia parece perder una propiedad importante y crucial: la accesibilidad inmediata a sus propios estados conscientes. Una vez que se desautoriza al propio sujeto para hablar acerca de sus estados conscientes, se allana el camino para dar un paso más: negar la fenomenología. Examinaremos ahora algunos casos de ilusiones ópticas en los que el componente fenomenológico se presenta como curiosamente persistente, y a partir de estos casos evaluaremos la postura de uno de los más famosos representantes de la postura negacionista con respecto a la fenomenología: Daniel Dennett.

⁴⁶ Conviene aclarar, sin embargo, que tiene sentido preguntar si realmente le duele en algún aspecto. Es una experiencia cotidiana del médico escéptico: “pero, ¿realmente le duele? Porque yo aquí no encuentro nada. A lo mejor le parece que le duele”. Es una afirmación legítima del médico. Lo cual demuestra que los enunciados de dolor tienen un uso objetivo y otro “fenomenológico”: “como quiera doctor, pero lo digo que me duele, y mucho”. Pero en este caso, lo que el médico quiere corroborar es si el padecimiento es psicósomático en virtud de que no encuentra los puentes que relacionarían ese dolor fenomenológico con una causa corporal. Sin embargo, si al paciente le parece₁ que le duele, entonces es cierto que le parece₁ que le duele: ese es el aspecto fenomenológico de su dolor y el médico no podría estar preguntando por *ese* aspecto. Lo que podría preguntar el médico es acerca de si su parecer₂ es correcto: quizás esté sintiendo un cosquilleo (ese sería su parecer₁) pero juzga que ese cosquilleo es un dolor (parecer₂). O, como ocurre en muchas experiencias posoperatorias: el paciente juzga que, si hace determinado movimiento, sentirá dolor. Entonces se abstiene de hacer ese movimiento. Pero en verdad no ha sentido el dolor, sino que infiere que tal o cual movimiento le causará dolor. Es más bien un temor al dolor (un parecer₂) que un dolor (un parecer₁)

Uno de los efectos de la ilusión óptica conocida como “the lilac chaser illusion” (también conocida como “el efecto pac-man”), ejemplificado en el modelo de la figura siguiente, consiste en provocar la impresión de un círculo color verde que se mueve alrededor de la cruz. (Figura 7)

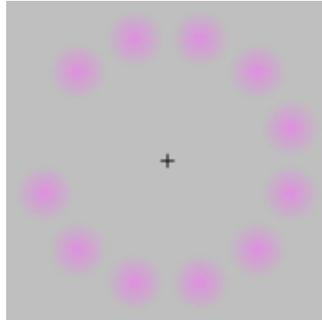


Figura 7. Usted está viendo un modelo estático de la figura que querríamos que viera. Dado que en un texto impreso resulta imposible lograr el efecto phi, lo remitimos a la dirección virtual <http://es.wikipedia.org/wiki/Archivo:Lilac-Chaser.gif>, donde está alojada una versión dinámica de este modelo

En el modelo original, los círculos de color lila se van apagando y encendiendo en secuencia. Si usted observa la cruz central durante treinta segundos, tendrá la impresión de que los círculos lila desaparecen y en su lugar sólo hay un círculo verde que gira alrededor de la cruz. El círculo verde aparece de modo incontestable en nuestra percepción; de hecho es tan incontestable que a veces no decimos que “nos parece” ver verde; simplemente decimos “*hay un círculo verde*”. Pero quienes saben que se trata de una ilusión óptica, pueden informar que “les parecía que había un círculo verde” girando alrededor de la cruz. En este uso, “me parece” equivale a “sin ninguna duda, percibo (tal o cual cosa)”, aun cuando el informante *sabe* que no hay un círculo verde; es decir, es un “me parece₁”

Cotejemos este uso del verbo ‘parecer’ con otro uso cotidiano, tal como lo distinguimos más arriba. Si hay un cartel en la panadería de la esquina y deseo leerlo a lo lejos, es posible que no alcance a distinguir las letras. Podría decir: “Me parece que dice que está cerrado por reformas”. Luego me acerco un poco más y descubro que había cometido un error: en realidad estaba cerrada por vacaciones. ¿Es este uso de “me parece” el mismo que el que utilizamos arriba con el fenómeno phi? En este segundo ejemplo, simplemente no vi las palabras del cartel, y “me parece” es una

inferencia a partir de lo que supuse, o lo que pude colegir de lo poco que vi. El “me parece”, en este caso, expresa una incertidumbre. En el ejemplo del círculo verde del efecto pacman, en cambio, el “me parece” nos da una información *segura* acerca de una percepción *inmediatamente accesible*. En el apartado anterior habíamos llamado “parecer₁” al estado de convicción inmediatamente accesible y “parecer₂” al parecer de segundo orden. Dentro de este parecer de segundo orden podemos incluir a las conclusiones de inferencias a partir de una información perceptiva deficiente.

Daniel Dennett en *La Conciencia Explicada* (pág. 373) propone un diálogo con un crítico que representa a “la tradición”, al que bautiza como “Otto”. Otto le señala el efecto de una ilusión óptica⁴⁷:

Dennett - En la contraportada, parece haber un anillo rosado ligeramente brillante.

Otto – No cabe duda de que es así.

D – Pues no hay ningún anillo rosado y brillante. No realmente.

O – De acuerdo. ¡Pero parece haberlo!

D – Efectivamente.

O - ¿Dónde está, pues?

D - ¿Dónde está qué?

O – El anillo rosado y brillante.

D – No hay ninguno; creía que usted acababa de reconocerlo.

O – Bueno, sí, no hay ningún anillo rosado ahí en la página, pero parece haber uno.

D – Efectivamente, parece haber un anillo rosado y brillante.

O – Hablemos, pues, de ese anillo.

D - ¿Cuál?

O – Ese que parece haber.

D – No hay ningún anillo rosado que simplemente parezca haber.

Da la impresión de que Otto no hace las objeciones correctas, y que el propio Dennett (más allá de la intención cómica del diálogo) es especialmente evasivo para responder. Traslademos el problema del “anillo rosado” al del fenómeno phi, el “círculo verde” que parece haber. Es cierto que en un sentido estricto “no hay ningún círculo verde” que parezca haber. Saber eso forma parte del conocimiento de lo que es una ilusión óptica. También es cierto que *parece₁* haber un círculo verde, y que podemos incluso describir la tonalidad del verde que parece haber, la posición y tamaño del círculo.

⁴⁷ La ilusión óptica a la que hace referencia forma parte de la contraportada de la edición en inglés, aunque no está presente en la edición en español.

Pero quizás Otto estaría mejor defendido si en lugar de preguntar por “el círculo rosa (o verde) que parece haber”, simplemente preguntara por “la sensación de percibir rosa” que está teniendo en ese momento. Su pregunta parece ir dirigida a la efectiva entidad de ese círculo rosa, y en ese sentido parece correcta la respuesta de Dennett: “No hay ningún círculo rosa”. No tiene sentido hablar acerca de una entidad *efectivamente existente* de un círculo rosa *inexistente* en la contraportada del libro. Pero sorprendería que Otto pusiera su énfasis en la presuposición de la existencia efectiva -e independiente de su percepción- del círculo rosa percibido. Otto debería estar empeñado en defender el aspecto fenomenológico *de su percepción* y no el hecho putativo de un círculo rosa *en la portada*. No tiene sentido preguntar “¿Dónde está” el círculo rosa, porque ambos saben que ese círculo no tiene existencia efectiva. En todo caso, es una mala pregunta que sólo sirve para que el oponente se luzca. No debemos culpar a Dennett de caricaturizar a su crítico: después de todo, el pequeño diálogo es casi un juego, más que la exposición de una polémica. Sin embargo, algunas de sus respuestas son discutibles: Otto le dice “hablemos pues de ese anillo” y Dennett, fingiendo no saber de qué está hablando, contesta “¿Cuál?”. La respuesta de Otto es “Ese que parece haber”. Dennett se niega a hablar de ese parecer. El énfasis de Otto está puesto en el *parecer* del anillo. Dennett descarta que se pueda hablar del anillo, porque no existe. Pero, ¿no puede hablarse acerca del *parecer* de la visión de un anillo rosa? ¿No puede hablarse del anillo *aparente*? ¿No puede compararse la apariencia de una ilusión óptica con otro tipo de apariencias? En el ejemplo que dimos más arriba, del cartel de la panadería que *parece* decir una cosa, pero en realidad decía otra, el fenómeno se puede explicar a partir de una visión deficiente (el cartel estaba muy lejos) y una inferencia a partir de lo poco que se había visto. No se diría, en ese caso, que se estaba *percibiendo inequívocamente* (*parecer₁*) al cartel como diciendo que estaba cerrado por reformas; en todo caso, a partir de la percepción de un cúmulo de letras apiñadas e indistinguibles, se indujo una interpretación que luego fue descartada con una mejor percepción.

¿Cuál es el “estado de percepción óptima” en lo que respecta a la ilusión óptica del anillo rosado o del círculo verde giratorio? No tiene sentido hacer este planteo; la propia ilusión óptica es un estado perceptivo pleno, inmediato, “claro y distinto”. No

es producto de una deficiencia de interpretación o de condiciones perceptivas: es una apariencia patente y no sujeta a los avatares inferenciales ni a caprichos subjetivos. Otto pregunta por ese estado perceptivo, y Dennett rechaza los términos de su pregunta, porque los encuentra mal planteados. Es verdad que así como lo plantea Otto suena insostenible. Lo que tal vez Dennett oculte con ese rechazo (y con esa caricatura de diálogo) es una afirmación más profunda: *es falso que usted esté viendo un círculo verde (o un anillo rosa), pero es cierto que le parece₂ que lo está viendo*. Pero, veamos: ¿Es falso que Otto ve el anillo rosa? ¿Es falso que usted ve un círculo verde en el fenómeno phi del ejemplo? (Figura 6) En realidad es *verdadero* que lo está viendo, y es *verdadero* que ahí no hay un anillo rosa / círculo verde. Usted ve algo que no existe, pero se puede hablar de eso que ve y que no existe, a menos que reservemos el verbo “ver” únicamente para los casos en los que efectivamente hay algo externo que sea causa de la visión (Esto lo discutiremos más adelante, en § 2.8 y § 2.9). Pero Dennett niega que se pueda hablar de esa percepción precisamente porque el contenido de la percepción no tiene contrapartida en un hecho externo a la propia conciencia. Su manera de abordar el diálogo parece implicar que, si no hay un hecho real involucrado, no tiene sentido hablar de la sensación “pura”, sin referencia a entidades que provoquen dicha sensación (el parecer₁). Saltaría más a la vista lo inadecuado (o lo evasivo) del planteo de Dennett si, en lugar de hablar de una ilusión óptica, habláramos del dolor en un miembro fantasma. El siguiente diálogo imaginario entre Otto y Dennett no provocaría la misma adhesión que el original, y sin embargo sigue su misma estructura:

Dennett – A usted parece dolerle el dedo índice de su mano izquierda amputada.

Otto – No cabe duda de que es así.

D – Pues usted no tiene mano izquierda. No realmente.

O – De acuerdo. ¡Pero parece dolerme!

D – Efectivamente.

O - ¿Dónde está, pues?

D - ¿Dónde está qué?

O – El dedo que me duele⁴⁸

D – No hay ninguno; creía que usted acababa de reconocerlo.

⁴⁸ Con este ejemplo salta más a la vista el absurdo de preguntar por “dónde está el anillo rosa” o “dónde está el dedo que me duele”. Ninguna persona bien informada, que estuviera objetando seriamente la posición de Dennett, haría esa pregunta. Es obvio que para Otto, ese dedo que duele no está presente en ningún sentido relevante para justificar su sensación de dolor.

- O – Bueno, sí, no hay ningún dedo porque no hay mano, pero parece dolerme.
 D – Efectivamente, parece dolerle.
 O – Hablemos, pues, de ese dedo.
 D - ¿Cuál?
 O – Ese que parece doler.
 D – No hay ningún dedo que simplemente parezca doler.

El doctor Dennett despacharía al paciente a su casa, pidiéndole que deje de referirse a entidades inexistentes. Evidentemente, sería un médico cruel e incomprensivo. Pero está claro que Otto no querría hablar *del dedo* en una circunstancia así, sino de *ese dolor* quimérico. Sin embargo la estrategia de Dennett consiste en negarse a hablar del dolor, dado que el objeto de dolor intencionado (el dedo que duele) no tiene existencia. Su actitud es congruente con su propia definición de sufrimiento: “El sufrimiento no consiste en verse atacado por un estado inefable intrínsecamente malo, sino en ver arruinados las propias esperanzas, planes y proyectos de vida por circunstancias impuestas sobre nuestros deseos, frustrando nuestras intenciones, sean estas cuales sean” (1995, pág. 461). De modo que un “dolor fantasma” sólo puede definirse como un tipo de sufrimiento si de algún modo arruina planes y proyectos, pero si exceptuamos estas consecuencias, no tiene sentido hablar del dolor en sí mismo. Para Dennett, sería correcto decir “Me parece₂ que tengo dolor en un dedo inexistente”, pero no: “Me parece₁ que tengo dolor en mi dedo de una mano que no tengo”. No hay una auténtica sensación (o, mejor dicho, la sensación no tiene el contenido intencional que el sujeto dolido refiere: sin duda, el dolor no puede estar en el dedo, porque en rigor no hay un dedo), sino una *apariencia* (un parecer₂) de sensación.

Hay un sentido (y es el que Otto y la persona con dolores en miembros fantasma deberían defender) en el cual *percibir* (ver, oír, oler, tocar, degustar) equivale a *parecer que se está percibiendo*. Si me parece que estoy viendo un círculo rosa (o un fantasma), es cierto que lo estoy percibiendo. En este sentido, que me parezca que percibo algo equivale a percibir ese algo⁴⁹. A continuación alguien podrá objetar: los

⁴⁹ Dejo de lado una posible objeción a esta ecuación: existen casos en los que el propio sujeto no sabe que está percibiendo lo que de hecho está percibiendo, o viceversa. Puede haber casos en los que el sujeto tenga dudas acerca de si siente dolor o si sólo es un cosquilleo molesto. En esos casos, podría

círculos rosa, al igual que los miembros fantasma (y que los fantasmas en general), no existen. Podemos dar por correcta esa afirmación, pero erra el punto. Si se tratara de fantasmas, en cierto sentido sería un alivio saber que no existen –aunque no lo sería descubrir que sufrimos de alucinaciones-, y en ese caso es importante distinguir si lo que percibimos “existe” o “no existe”. En el caso del anillo rosa o del miembro amputado, ya *sabemos de antemano* que no existe. Sin embargo, podemos hablar detalladamente acerca del anillo rosa: podemos incluso, si se nos dan muestras adecuadas, decidir cuál es el tono de rosa que hemos visto o incluso la posición y el tamaño del anillo visto. Quienes tienen miembros amputados pueden, en algunos casos, detallar incluso la posición en la que sienten que está su mano y sus dedos inexistentes. Es decir, no sólo podremos hablar de percepciones que son producto de cualidades secundarias ilusorias, sino también de percepciones ilusorias de cualidades primarias. Cuando Otto dice “Hablemos pues (de ese anillo, de ese dedo)”, Dennett finge que no hay nada de qué hablar. En realidad hay algo de qué hablar: la apariencia específica y detallada de esas sensaciones (el parecer₁); lo que no hay es el objeto intencionado, y eso lo saben de antemano ambos interlocutores.

La estrategia de Dennett consiste en matar a la fenomenología negándose a hablar de ella. Todo lo que hay es texto que se añade al texto; que se revisa y se actualiza. Pero no hay una “apariencia real” de las cosas. No hay una “apariencia real” del anillo rosa, ni del círculo verde, ni del dolor:

“(…) Parece haber fenomenología. Este es el hecho que, con entusiasmo, admite el heterofenomenólogo. Pero de este hecho innegable y universalmente admitido no se deduce que realmente haya fenomenología.” (1995, pág. 377)

El heterofenomenólogo es el partidario de la heterofenomenología, esto es, quien examina las autoatribuciones fenomenológicas que realizan los sujetos (mediante juicios fenoménicos) y las somete a estudios controlados para contrastarlas.

Una página antes:

(Otto pregunta) «Pero, ¿qué me dice de la fenomenología real?»

decir “me parece₂ que me duele”, usando el verbo “parecer” de un modo cauteloso, evitando la afirmación segura del dolor. Pero en lo que respecta al ejemplo del anillo rosa (o del círculo verde), la ilusión óptica es tan contundente e innegable que no deja lugar a una utilización del “parecer” en este sentido.

(Dennett responde) - No existe. (1995, pág. 376)

¿A qué viene aquí la pregunta por la fenomenología *real*? La pregunta de Otto se parece a la de los interlocutores socráticos de algunos diálogos platónicos: sus preguntas o comentarios sólo son estrategias retóricas para que Sócrates (es decir, Platón) desarrolle su tesis sin contratiempos. ¿Por qué Otto haría esa pregunta? O, mejor dicho, ¿En qué está pensando Dennett cuando inventa a un crítico que le hace esa pregunta?

La pregunta por la “apariencia real” de un hecho es desconcertante. Presupondría, en principio, que hay apariencias reales y apariencias irreales, y que a cada tipo de apariencias les corresponde una fenomenología (fenomenología “real” e “irreal”). Si hay algún modo inteligible de entender esta distinción fenomenológica, podríamos encontrar algunos ejemplos en los que nos “parece” que algo tiene una “apariencia real”, y otras veces nos “parece” que algo tiene una “apariencia irreal”. No debemos identificar, sin embargo, la “fenomenología irreal” con una “apariencia alucinatoria”, mientras que la “fenomenología real” sería el contenido mental de una “apariencia no alucinatoria”. Si este fuera el tipo de identidad en el que piensa Dennett, entonces la pregunta de Otto, acerca de la fenomenología *real* de una experiencia *ilusoria* no tendría ningún sentido. Parece que, aun en experiencias ilusorias, Otto se cree con derecho a preguntar por la *apariencia real* de ese hecho ilusorio. Una cosa es afirmar, como hace Dennett, que la *apariencia* no es un hecho mental complejo ubicado en alguna zona específica del cerebro. Otto no está afirmando (al menos no debería hacerlo si fuera un crítico inteligente y no una dócil caricatura retórica) que el anillo rosa está “pintado” con pigmento en su mente⁵⁰, como tampoco afirma que el anillo rosa está ahí, en la contraportada del libro. Podemos prescindir de la noción de pigmento y de significador central, y sin embargo todavía hay espacio para preguntar por una fenomenología, esto es, por el modo en que se dan las experiencias conscientes. Otra cosa, en cambio, es negar que haya un *modo de darse* de las experiencias, incluidas las ilusorias, como parece hacerlo Dennett al negar de plano la

⁵⁰ “Presumo que nadie piensa realmente que la «repleción» sea un proceso en que el cerebro se dedica a cubrir un espacio con un *pigmento* (...) Parece que, efectivamente, hay algo en todos estos casos [de repleción], algo que el cerebro tiene que proporcionar (para «rellenar»). ¿Cómo podríamos llamar a esta cosa desconocida? Llamémosle *figmento* ” (Dennett, 1995, pág. 357)

fenomenología. “Pero parece haberla”, replica Otto. “Exacto, parece haber fenomenología.”, contesta Dennett (1995: 377). Pero, ¿es necesario que Otto indague acerca de la fenomenología? Otto ha replicado que “le parece” que hay un anillo rosa, y esa sola apariencia es un hecho fenomenológico. Pero del hecho de que me parezca que hay un anillo rosa o que me parezca que me duela no se deduce que me parece que hay fenomenología. Corremos el riesgo de que el lenguaje nos tienda una trampa: la afirmación de que el universo sólo es un conjunto de apariencias no implica que el juicio fenomenológico de apariencia sea, a su vez, *aparente*. Si la fenomenología es aparente, entonces podemos preguntar por una fenomenología de la fenomenología (“¿Cómo es que me parece que hay fenomenología?”), y de este modo no nos liberamos de la pregunta fenomenológica, sino que nos embarcamos en un indeseado *more metafenomenológico*. Una vez más, ¿por qué Otto haría esa pregunta acerca de la fenomenología? Todo lo que él necesita es una explicación de *ese anillo rosa* que ve y del que *sabe* que no existe. Esa explicación, que le es negada, es toda la fenomenología que hace falta en este caso.

Según Dennett, la fenomenología parecería necesaria si suponemos que la conciencia está *llena* y si creemos en un significador central o teatro cartesiano. En cualquiera de estos casos, se supone que el sujeto es unitario (es decir, su conciencia, su juicio y su aparato perceptivo trabajan de forma seriada, ordenada y unívoca) con respecto a sus propias percepciones y que puede declarar sin error lo que está “percibiendo en este momento”. Dado que ninguna de las condiciones se satisface (no hay un hecho único de conciencia sino un conjunto continuamente actualizado de borradores; el juicio es corregible en toda instancia de la percepción e incluso con respecto a percepciones pasadas, y la estimulación o la idiosincrasia del aparato perceptivo no determina de forma unívoca el juicio que se emite a partir de ese hecho de conciencia no centrado⁵¹) un sujeto no puede juzgar de manera inmediata e indubitable si realmente tiene o no tiene una sensación determinada. Puesto que, tal como nos lo muestra Dennett, no existe este sujeto cartesiano para quien se monta un teatro completo y coloreado con figmento, la propia noción de fenomenología parece desvanecerse. Después de todo,

⁵¹ “No centrado”: en dos sentidos: 1) no existe un lugar central donde se “da” el hecho consciente acabado; 2) lo que llamamos “hecho consciente” es una verbalización parcialmente arbitraria de un estado de cierto conjunto de borradores. No hay una “versión definitiva” de un hecho consciente.

hablar acerca de un estado consciente no es más que seleccionar de un modo parcialmente arbitrario uno de los múltiples borradores de la mente y erigirlo como el representante de “lo que verdaderamente me está / estuvo ocurriendo”: “Si queremos fijar un instante del procesamiento en el cerebro como el instante de la conciencia, la decisión será arbitraria” (1995: 139); “(...) [E]n cualquier intervalo de tiempo hay múltiples versiones de fragmentos narrativos en varios estados de revisión, en varios puntos del cerebro” (1995: 149) Pero aun si aceptamos la noción de borradores múltiples y si negamos tanto el significador central como el teatro cartesiano, ¿tenemos que aceptar, *ipso facto*, que *cualquier* selección de borradores sea (parcialmente) arbitraria? Piénsese una vez más en el diálogo con Otto. Cuando Otto insiste en una explicación de la visión de ese anillo de color rosa, ¿podemos decir que sus demonios (las entidades caóticas responsables de sus emisiones lingüísticas) han elegido simplemente “un ganador por defecto que acaba siendo proferido” (1995, pág. 251) a partir de un “(...) proceso que avanza por múltiples vías, se ejecuta en unos cientos de milisegundos, durante los cuales se producen diversos añadidos, incorporaciones, enmiendas y sobrescrituras de contenido, en diferentes órdenes [los cuales c]on el tiempo (...) dan lugar a algo *bastante parecido* a un flujo o secuencia narrativa, que podemos concebir como sujeto a revisiones continuas por parte de diversos procesos distribuidos por el cerebro, y que se prolonga indefinidamente en el futuro”? (1995, pág. 148). Si esta fuera la pintura correcta de lo que ocurre cuando informamos acerca de lo que *nos parece*, entonces debemos coincidir en que nunca podremos dar una información fidedigna. No sólo no tenemos acceso privilegiado a lo que “nos parece”, sino que ese “parecer” es apenas un suceso rescatado de la continua marea, siempre renovada, siempre en movimiento, de los procesos mentales. ¡Porque quizás no nos parecía lo que dijimos que nos parecía! O, en un sentido más estricto, no hay –desde el punto de vista científico– cosas tales como los pareceres, los aspectos cualitativos de las sensaciones o las ilusiones ópticas. Estos tres hechos mentales son sólo *apariencias* de hechos mentales. Resulta extremadamente difícil desembarazarse del lenguaje de lo aparente, y da la impresión de que vuelve a surgir en cuanto se lo intenta eliminar. Sin embargo, en cuanto el juicio fenomenológico levanta su cabeza, Dennett intenta ubicarlo en su lugar: las afirmaciones que un sujeto hace acerca de sus propios estados de conciencia deben ser sometidas a investigación

empírica, y la investigación arrojará, acerca de su propio parecer, resultados diferentes de lo que él mismo enuncia que le parece. El juicio de parecer en primera persona está, siempre, sujeto a investigación. No es ilegítimo, pero no tiene la palabra autorizada. Es, en todo caso, una opinión más: usted creía que le había parecido que había un anillo rosa, pero en realidad no había un anillo rosa, de modo que la creencia acerca de su parecer, si bien es plausible desde su punto de vista, es errónea desde el punto de vista heterofenomenológico. No puede parecerle que hay un anillo rosa si no hay fenomenología en absoluto; que a usted le parezca que hay un anillo rosa es algún tipo de *flatus conscientiae*... Aun cuando a usted no *le parece* que lo sea. Una vez más (como más arriba se señaló), el *parecer* reaparece, pero sometido a una cascada de meta-pareceres que lo convierten en un hecho mental un tanto inasible y poco familiar. Como en la pesadilla epistemológica, el parecer de un nivel genera un parecer en el metanivel: ¡No parece, pero parece parecerme! Las oraciones del tipo “me parece que percibo X” están en un curioso estado de relaciones veritativas con respecto a su oración subordinada: aun cuando sólo intentan informar un suceso fenomenológico –un parecer₁– sin someterlas a un juicio acerca de su valor de verdad, en realidad son siempre falsas. El sujeto *nunca* está en posición de comunicar lo que *él considera* como contenidos inmediatos de su percepción.

Michael Tye (1993) encuentra que en ciertos pasajes de *La conciencia explicada* (algunos de los cuales hemos reproducido más arriba) Dennett identifica el parecer con los juicios acerca de características externas. Dado que el parecer se ejecuta mediante un acto lingüístico (un juicio, una decisión o una opinión), Dennett asume que tal acto lingüístico y el objeto de la apariencia son idénticos. Tye propone un contraejemplo para mostrar que no hay identidad entre una creencia y una apariencia:

Upon seeing wheels rotating rapidly enough in one direction, it may be well look to me that they are rotating rather slowly in the opposite direction. But I don't believe this. For I am familiar with the illusion. Perhaps Dennett will respond that we should focus only on judgments formed without interference from other judgments about the seen object (assuming there is one). This does not help, however, since there may not be any other relevant beliefs that interfere. Suppose, for example, I am seeing the Muller-Lyer figure for the first time, and I am disposed to be very cautious about what I believe. I may well not believe that the two lines are really of different lengths even though that is how they appear. (1993, pág. 896)

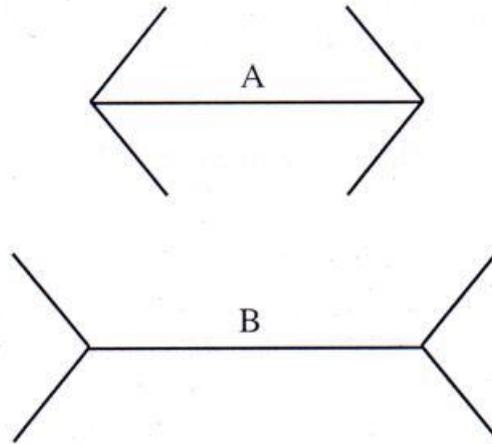


Figura 8. Figura de Müller – Lyer.

Si ya conocemos cómo funciona una determinada ilusión óptica, o si sospechamos que podemos estar siendo víctimas de una, nuestra actitud es la de separar la convicción inmediata (el parecer₁) de la creencia (el parecer₂). Pero cuando Otto pide hablar de ese anillo, parece que se encontrara imposibilitado de hacer tal separación. Su actitud es la de una persona que *sigue creyendo* en la ilusión y no puede separar su parecer inmediato de la creencia de que algo no anda bien con ese parecer. Remarquemos una vez más que esta no es la respuesta usual y esperable ante una ilusión óptica.

Por otra parte, es posible que algo parezca ser azul sin que nos formemos un juicio, como presumiblemente ocurre en los niños pequeños:

Surely something can seem to one to be blue, say, without one having the capacity to form a belief with an intentional content as complex as the view demands, a belief that it seems to one that something is blue. Consider, for example, the case of young children (1993, pág. 896)

Algunos capítulos antes de incluir el diálogo con Otto (el cual es, visiblemente, una exageración informal y en cierto modo paródica), Dennett imagina una objeción de parte de un sujeto que se somete a un experimento del fenómeno phi de los colores (1995, pág. 147). Aunque el experimento no es el mismo, la objeción vale también para el que hemos puesto aquí como ejemplo, cambiando “punto” por “círculo verde” (Véase Figura 7):

Sé que no había realmente un punto en movimiento en el mundo – después de todo, no es más que movimiento aparente-, pero también sé que me pareció que el punto se movía, así que además de mi juicio de que el punto pareció moverse, existe el evento de que trata mi juicio: el hecho de que el punto pareció moverse. No hubo movimiento real, así que debe haber una apariencia real de que el punto se movió sobre la cual se basa mi juicio. (1995, pág. 147)

Dennett argumenta, contra esta objeción, que “El postular una «apariencia real» además del juzgar o el «considerar» expresado en el testimonio del sujeto es multiplicar las entidades sin necesidad.” (1995: 147)

Consideremos por separado la objeción y la respuesta de Dennett. El sujeto que realiza la objeción (tal vez un Otto más refinado) admite que el círculo en movimiento no estaba “en el mundo”. Esto es un paso importante que Otto no había dado cuando preguntaba “dónde está el anillo rosado”. Sin embargo, este Otto refinado hace una concesión demasiado aventurada: hay una “apariencia real” de un suceso. Ya hemos dicho algunas palabras sobre ciertas consecuencias desconcertantes de esta invocación a una “fenomenología real” (página 150). Si alguien postula la existencia de apariencias “reales”, también postularía apariencias “irreales”. Pero no queda claro dónde trazar la línea, pues el círculo verde / anillo rosa son apariencias “reales” de hechos ilusorios (y por lo tanto “irreales”). Sin embargo, quizás no haga falta buscar la línea demarcatoria entre las apariencias reales y las irreales. Quizás hay algo en el uso de estas nociones que es incorrecto.

Existe una construcción de lenguaje cotidiano sorprendentemente parecida a la de “apariencia real”, tanto en su forma como en la perplejidad que genera. Es la noción de “posibilidad real”. Se supone que ciertas posibilidades (por ejemplo, que esta noche cene carne asada) son “reales” mientras que otras serían “irreales” o, en todo caso, (si la matización fuese aceptable) tendrían menor grado de realidad. Un ejemplo de este último caso sería la suposición de que esta noche yo ganara un auto Ferrari cero kilómetro. ¿Puede trazarse la línea que divide a las posibilidades reales de las posibilidades irreales? (Dejemos de lado, por supuesto, las manifiestas *imposibilidades*, que no deberían ser confundidas con las posibilidades irreales. Si la noción de “posibilidad irreal” tiene algún sentido, en principio se supone que no es el mismo que el de “imposibilidad”) Si las condiciones actuales no nos permiten visualizar

determinadas posibilidades como actualizables en lo inmediato, y dado nuestro conocimiento de las condiciones actuales, entonces parece justificado decir que cierta posibilidad es irreal. Las personas que construyen castillos en el aire suelen barajar posibilidades irreales, o por lo menos fácticamente implausibles. Sin embargo, *podría* ocurrir que esta noche yo gane una Ferrari cero kilómetro, y ese suceso sería tan real como el hecho de cocinar carne asada. En un estado de cosas en el que ambos sucesos ocurren, ambas posibilidades son (o fueron) reales. ¿Cómo distinguimos, entonces, las posibilidades reales de las irreales? Quizás hay algo que no anda bien con los calificativos *real* e *irreal* aplicados a la noción de posibilidad. Una posibilidad se define, precisamente, por su *irrealidad* actual, aunque es una proyección de una *realidad* alternativa futura. ¿Qué le añaden los calificativos de “real” e “irreal” a la palabra “posibilidad”? Lo único “real” que hay en una “posibilidad real” es, precisamente, la *realidad* alternativa futura. Y, en su contrapartida, lo único irreal que hay en una posibilidad es *la no realización actual* de dicha posibilidad. Ahora bien, si con “posibilidad real” se está queriendo decir “una posibilidad que tiene muchas chances de ser real”, en contraposición con otras posibilidades que tienen baja chance, entonces los calificativos utilizados son incorrectos. Deberíamos hablar de posibilidades esperables, posibilidades por las que vale la pena apostar, en lugar de posibilidades “reales”. En todo caso, tal como afirma Wittgenstein (§ 6.4) acerca de las proposiciones, todas las posibilidades “valen lo mismo” (1999, pág. 177)

Quizás este mismo hechizo del calificativo “real” esté afectando, también, al término “apariencia”. Una apariencia es, de acuerdo a su propia definición, un determinado aspecto de algo que se aparece a la conciencia. La apariencia es, por ello, algo sobre lo que no puede juzgarse como real. Una apariencia real es, en todo caso, algo de lo cual ya se ha hecho un juicio, y de ese modo, o bien no es ya una apariencia, o bien debemos negarle el calificativo de “real”. En contrapartida, decir que una apariencia es “irreal” equivale casi a afirmar el propio carácter de la apariencia (la apariencia no es lo real: es, en todo caso, un tipo de suceso consciente con algún tipo de contenido, del cual no se afirma que sea “real”). Pero todas las afirmaciones de apariencia “valen lo mismo” en lo que respecta al juicio sobre su realidad o su irrealidad: el calificativo no añade una nueva distinción.

De modo que la objeción, así planteada, puede rechazarse. Resulta fácil argumentar contra una objeción que utiliza un andamiaje conceptual discutible. Por eso, cuando Dennett despacha rápidamente la objeción, desdeñando la noción de “apariencia real”, está en lo cierto. No tiene sentido multiplicar las entidades: además de la “apariencia” lisa y llana, que puede identificarse con el juicio expresado en el testimonio del sujeto, no hay “apariencias reales” y “apariencias irreales”, pues tales distinciones parecen violentar la navaja de Ockham. Nada que objetar a esa argumentación. Pero, precisamente porque la objeción está mal planteada, nos distrae de la *verdadera* objeción que el Otto refinado debería haber interpuesto: lo que hacía falta plantear, en lugar de enredarse con las “apariencias reales” era que la *apariencia* puede separarse del *juicio*. Pero Dennett no nos presenta las cosas de ese modo; mediante la argucia de un objetor que introduce conceptos discutibles, nos distrae y se dispensa de explicarnos por qué el juicio es todo lo que se necesita para explicar el parecer (¡Y por qué *parece* que no es así!)

Ahora bien, ¿qué tipo de explicación esperarían el Otto inicial y el Otto refinado? Ambos desean que se les explique su *parecer*. Ambos tienen la impresión de que algo queda inexplicado o, mejor dicho, reducido al concepto de *juicio*. Todo lo que usted pide que se le explique, argumenta Dennett, puede ser perfectamente explicable en términos de una teoría de los juicios. La heterofenomenología, que se basa en los juicios en primera persona, no necesita postular una misteriosa entidad interna (el *quale*, la entidad fenomenológica básica) que sirve de referencia a esos juicios. “El hecho de eliminar algo no es un rasgo de las explicaciones fallidas, sino de las explicaciones logradas” (1995, pág. 465). Otto tiene la sensación de que ha sido estafado. Él esperaba que le explicaran qué es este anillo rosado o ese punto verde en movimiento, pero todo lo que recibe es la reprimenda, el recordatorio de que no hay un anillo o un círculo que se mueve. De manera más radical, se le dice que no hay fenomenología. Si todavía exige una explicación, todo lo que puede hacerse es hablar en términos de sus estructuras perceptuales, la reflectancia de la luz, el modo en que su cerebro resuelve determinados patrones luminosos (es decir, una explicación en términos de óptica y de fisiología), pero su fuerte *parecer* de un anillo rosa en la portada no se explica en absoluto mediante la reducción de su fenomenología a

procesos ópticos y fisiológicos. O, al menos, él no se considera satisfecho con esa explicación. Su experiencia consciente, las cualidades específicas del rosado y de la esfericidad del anillo requieren de una explicación ulterior, un salto por encima de la fisiología y de la óptica. ¿Pero no estará Otto malinterpretando la naturaleza de las explicaciones? Dennett (1995, pág. 466) nos recuerda:

Los sólidos, los líquidos y los gases pueden explicarse a partir de cosas que no son sólidos, ni líquidos ni gases. Sin duda la vida puede explicarse a partir de cosas que no están vivas, y esa explicación no deja sin vida a las cosas vivas. La ilusión de que la conciencia es la excepción proviene, sospecho, del hecho generalizado de no haber sabido comprender este rasgo general de la explicación. Al pensar, sin razón, que la explicación comporta suprimir algo, creemos salvar aquello que en caso contrario se perdería, volviéndolo a colocar en el observador como un *quale* o como una propiedad «intrínsecamente» maravillosa.

De modo que quizás Otto (y el lector desprevenido) le están pidiendo a las explicaciones algo que las explicaciones científicas no pueden dar. El punto, quizás, es decidir si de hecho una explicación como la que propone Dennett es realmente una explicación (al menos parcial) de la conciencia o una explicación de los dispositivos cerebrales y perceptuales para que haya conciencia. Puede ser discutible que toda explicación necesite reconceptualizar el hecho explicado de manera tal que ese mismo hecho desaparezca. Un ejemplo de este tipo es el de las explicaciones fundamentales, en las cuales se asume que el hecho a explicar no puede reducirse a otros hechos (Chalmers, 1999), o las explicaciones históricas tal como las proponía Locke (v. § 1.7). También, si bien puede aceptarse que cierta clase de explicaciones no tiene por qué dar cuenta de hechos locales en grano fino, eso no implica que tal grano fino no esté presente. La explicación puede ser reductiva, pero el hecho explicado no tiene por qué ser *eliminado* como consecuencia de la reducción operada (del mismo modo como no se eliminan los objetos macroscópicos a partir de una explicación por sus componentes microscópicos). El parecer no es una entidad y por lo tanto nadie que pidiera una explicación del parecer estaría pidiendo una ontología del parecer. Por eso resulta necesario indagar si la eliminación del parecer en la explicación es un procedimiento reduccionista legítimo. En rigor no hay una *entidad* a ser reducida, de modo que tal vez la reducción operada sea inatinerente. Por otra parte, los juicios de parecer se resisten a ser eliminados.

¿Qué podemos hacer con el hambre epistémico de Otto acerca de la visión de ese inexistente anillo rosa?. Quizás Otto se sentiría aliviado si, en lugar de negársele insistentemente la existencia de ese anillo, se le hablara de los distintos tipos de ilusiones visuales.

Existen tres tipos de ilusiones visuales: ilusiones ópticas, ilusiones sensoriales e ilusiones perceptivas. John Eccles (Eccles, John & Zeier, Hans, 1985), siguiendo a Gregory (Gregory, 1968) establece que:

Las ilusiones ópticas tienen una explicación puramente física, como sucede con los espejos, prismas y lentes y sus equivalentes en la naturaleza; así, por ejemplo, la desviación aparente que sufre una barra en la superficie de contacto entre el aire y el agua, y los espejismos. Las ilusiones sensoriales son atribuibles a procesos fisiológicos en la retina y el cerebro. De este tipo son, por ejemplo, las imágenes persistentes y las percepciones originadas por espirales giratorias. (...) Todas las ilusiones pertenecientes a estos dos tipos se explican satisfactoriamente por procesos que ocurren en el mundo 1, sea en el mundo externo o en los órganos receptores y el cerebro del observador. Otra cosa ocurre con las ilusiones perceptivas, que dependen del juicio consciente del observador y pertenecen, por tanto, al mundo 2. Las más simples son las conocidas figuras de profundidad ambigua, como el cubo de Necker o la escalera de Schroeder que puede interpretarse como una escalera ascendente o como una cornisa saliente. (Eccles & Zeier, 1985, pág. 145)

En este pasaje Eccles intenta reforzar la peculiar idea de que las ilusiones perceptivas, en las que una figura se interpreta de dos modos alternativos, no son producto de una transformación global en los patrones cerebrales sino de una interpretación que realiza directamente la mente, por principio separada del cerebro aunque conectada con él mediante patrones de información. Por ello, las ilusiones perceptivas son ilusiones de la propia mente que no poseen contrapartida fisiológica. El artículo de Gregory, de donde extrae la clasificación de las ilusiones perceptivas, se abstiene de esta conclusión acerca de la naturaleza de la mente. La división entre mundos 1 y 2 corresponde a la división de Popper entre el mundo físico (mundo 1), mundo mental (mundo 2) y mundo conceptual (mundo 3) (Popper, 1986, págs. 136-7)

Según el pasaje citado, el tipo de ilusión que afecta a Otto es del tipo 2: una ilusión sensorial, esto es, analizable a partir de procesos fisiológicos en la retina y el cerebro. Se le mostraría que, de no ser por la conformación de su retina y del modo en que su cerebro percibe el movimiento, la ilusión del anillo rosado (o del círculo verde) sería

imposible. O, tal vez, que una explicación de los procesos fisiológicos de la retina y del cerebro es suficiente para provocar la comprensión del fenómeno. El fenómeno se convierte, entonces, en una *consecuencia necesaria* de la actuación de determinados mecanismos visuales y cerebrales. Si en lugar de que Dennett le señalara que “el anillo no existe”, le ofreciera con detalle la explicación antedicha, ¿se sentirá Otto satisfecho con ella?

La Tradición que representa Otto asume (tal como Locke lo había hecho) que hay una desproporción entre el aspecto cualitativo de la conciencia y los hechos fisiológicos que los explican. En otras palabras, que queda resto explicativo, de modo que la explicación anterior sólo explica los procesos causales en el mundo 1 para generar un fenómeno cualitativo en el mundo 2, aunque no explica en sí misma el suceso del mundo 2. Pero la explicación que proporciona Dennett asume que no existe una división de mundos, pues el *juicio* fenoménico (a lo que se reduce la experiencia cualitativa) está presente en el mundo 1 y es por lo tanto heterofenomenológicamente accesible. Nuevamente, ¿es acertada esta ecuación entre juicio y experiencia cualitativa? Asier Arias Domínguez (2011) expone los problemas que surgen de las relaciones entre juicio y experiencia:

Ciertamente las relaciones entre estos extremos (juicio o creencia y experiencia o qualia) resultan de lo más peliagudas (a nuestro parecer por su estrecha vinculación con el tratamiento de delicados problemas relacionados con el papel que el lenguaje y el uso de conceptos juegan en la conciencia). Señalemos, no obstante, que Dennett [en *La Conciencia Explicada*] trata de reducir experiencia a juicio (cosa que se ve particularmente bien en su método heterofenomenológico), pero, cabe aducir, una creencia sobre una experiencia no equivale a una experiencia consciente ni es idéntica a ella – al menos no necesariamente; el matiz es importante, porque cabe defender que tanto las creencias como el resto de las actitudes proposicionales, que a menudo son tratadas como estados mentales portadores exclusivamente de contenido intencional y carentes tanto de aspecto fenoménico, pueden ser experimentadas conscientemente de diferentes maneras. (2011, pág. 19n)

Collin Mc Ginn (1995) declara que Dennett no hace un esfuerzo serio para mostrar cómo las características fenoménicas de los estados conscientes pueden ser eliminadas, aun cuando se mantenga la apariencia de que siguen existiendo. Con respecto a la ausencia de las propiedades fenoménicas, Mc Ginn responde:

For on the crucial question, long deferred by Dennett, of whether he really believes in such phenomenal properties as the smell of a rose, we get a very puzzling answer: that

there only *seem* to be such seeming, but really there aren't any. He writes: 'There seems to be phenomenology'. That's a fact the heterophenomenologist enthusiastically concedes. But it does not follow from this undeniable, universally attested fact that there really is phenomenology' (p.366) Well, I would have thought that it did follow, at least if Dennett is using language normally, since it is self-contradictory to say that there are no seemings but that it seems there are. That's like saying there are no feelings such as the feeling of pain but there are feeling *of* feelings of pain, or that there are no beliefs just beliefs in beliefs. (...) Anyway, there surely is phenomenology, if all we mean by that is that people have experiences. I suspect that Dennett denies even this, but he remains obscure on this simple question (Mc Ginn, 1995, pág. 247)

Dennett parece cometer una petición de principio al aceptar el *parecer parecer*, pero no el *parecer* a secas. Está haciendo una analogía demasiado aventurada: así como el anillo rosa parece real pero no lo es, toda la fenomenología parece "real", pero no lo es. Da la impresión de que desea negar los modos de darse de un determinado estado mental. Si Dennett negara que cada instanciación de una creencia (por ejemplo: la creencia en este preciso momento de que los gatos son demonios, y *la misma creencia* expresada ayer a la tarde) tiene cada vez un contenido diferente *para el individuo que instancia esa creencia*, o un *modo de darse* distinto, entonces caería en la implausible tesis de que todos los actos de creencia acerca de un hecho son idénticos para el mismo individuo en todos los momentos en que ejecuta esa creencia conscientemente, y son idénticos para todos los individuos. Esta es una consecuencia del funcionalismo eliminativista: si los estados cualitativos se definen funcionalmente, entonces las diferentes implementaciones de un mismo estado cualitativo no pueden diferir. "Cómo es ser un murciélago" no es diferente a "Cómo es ser un murciélago robot", si ambos (el murciélago y el murciélago robot) encarnan la misma estructura funcional, entonces ambos se "sienten" del mismo modo (aun cuando les parezca que no se sienten del mismo modo, si hubiera algún medio de 'comparar los pareceres'). De manera similar, el parecer ver rojo hoy y el parecer ver rojo ayer no tienen diferente contenido, dado que en ambos casos se actualiza una misma estructura funcional. Steven Horst (Horst, 2007) caracteriza a esta versión del funcionalismo del siguiente modo:

Functionalists view mental state types (such as "pain") as functional types, and point out that they, like other functional types, are multiply realizable. Assuming that "pain" is a functional-kind term, humans, earthworms, Martians, and even robots might have

functional states corresponding to pain, yet be built out of very different types of components, so that the “realizer” of the pain-function in each is unique. Nonetheless, human pain, earthworm pain, Martian pain, and even robotic pain could each be physical processes, in the sense that in each case the pain-function is realized entirely by some physical system or other. More radically, each instance of my having an occurrent belief with a given content (say, each time I think “Mark Twain wrote Tom Sawyer”) might be realized by a slightly different pattern of neural activity in my brain; but each occurrence that realizes that belief-that-p function is some physical state or other” (2007, pág. 99)

Si se niega la fenomenología, los estados cualitativos pueden definirse funcionalmente, sin tener en cuenta el tipo de implementación física ni el tipo de instanciación de ese estado: un robot, un murciélago o un marciano estarían hablando de *la misma sensación* cada vez que utilizan el término “dolor”, aun cuando su dolor estuviera implementado en estructuras físicas radicalmente diferentes. Pero no tendría sentido que cada uno se refiriera a “el dolor que tengo yo” (en tanto individuo de esta especie, en estas circunstancias, en estas condiciones epistémicas) como algo que merezca una atención especial independientemente del concepto funcional de dolor.

Otto sólo puede “señalar su sensación” de anillo rosado, y cuando pide una explicación para esa sensación, le estaríamos explicando las propiedades funcionales de su propio sistema perceptivo que provocan las ilusiones sensoriales. Pero Otto podría insistir, aun, en que no es eso lo que pide que le expliquen. Él asume que, aunque le parece correcta y aceptable esa explicación, todo lo que se ha hecho es correlacionar su estado de visión ilusoria (mundo 2) con un estado de su sistema perceptivo (mundo 1); que puede entender que las propiedades de su sistema perceptivo más un estímulo adecuado son necesarias para causar esa ilusión, pero que aun así no son suficientes. Hay un resto que queda inexplicado, y ese resto es, precisamente, el hecho fenoménico del anillo rosa. Locke, sin duda, podría suscribir a la objeción de Otto. Tanto Locke como Otto dirían que su visión del anillo rosa no ha quedado explicada en absoluto mediante la apelación a mecanismos funcionales. Dennett da por suficiente una explicación relacionada con los mecanismos del mundo 1, eximiéndose de dar una ulterior explicación del anillo rosado que ve Otto. ¿Quién de los dos está en lo correcto con respecto a la explicación del hecho?

Otto (y con él la tradición, uno de cuyos representantes es Locke) pide más contenido explicativo; Dennett descompone el fenómeno y dice “esto es todo lo que hay que explicar”. Uno considera que el explanandum queda sin explicar. El otro, muestra que el propio explanandum fenomenológico no tiene lugar *qua* fenomenología. Ambos tienen expectativas diferentes acerca de lo que cuenta como explicación.

Quizás una parte de las expectativas de Otto pueda resumirse en las demandas de Nagel expuestas en su famoso artículo, *Cómo es ser murciélago* (1983). Al final del artículo, Nagel supone que:

[E]s posible asomarse a la brecha que media entre lo subjetivo y lo objetivo desde otra dirección. Dejando a un lado por el momento la relación entre la mente y el cerebro, podemos buscar una comprensión más objetiva de lo mental por derecho propio. En este momento carecemos de todo medio para pensar en el carácter subjetivo de la experiencia sin apoyarse en la imaginación, sin adoptar el punto de vista del sujeto experiencial. Cabría considerar esto como un desafío a la elaboración de nuevos conceptos y a la adquisición de un nuevo método, una *fenomenología objetiva* que no dependa de la empatía ni de la imaginación. (Nagel, 1983, pág. 520 cursivas añadidas)

La sugerencia de una “fenomenología objetiva” parece caer en el mismo problema que veíamos con respecto a “posibilidad real” o “apariencia real”. Una fenomenología objetiva ya hace un juicio acerca de la naturaleza de lo fenoménico: si es objetivo, puede ser describible no sólo como un fenómeno, sino también como un acaecimiento. Pero, ¿en qué se distinguiría una fenomenología objetiva de una heterofenomenología? Nagel agrega: “Los rasgos estructurales de la percepción podrían ser más accesibles a una descripción objetiva, aun cuando algo quedase excluido” (pág. 520) En principio, debería ser posible explicar a alguien que nunca vio el color rojo, *cómo es ver rojo*, aunque esto no le haga tener la experiencia de rojo (algo habrá de quedar excluido)

En una reflexión sobre el texto de Nagel, D. Hofstadter (Dennett, Daniel; Hofstadter, Douglas, 1983) encuentra que la especulación de Nagel no tiene sustento. Lo que Nagel pide que se le explique en rigor no puede satisfacerse. Las explicaciones no llegan hasta allí o, mejor dicho, no hay nada que necesite una explicación:

Este es el punto urticante del artículo de Nagel. Quiere saber si es posible ofrecer, en sus propias palabras, una descripción (de la verdadera naturaleza de la experiencia

humana) en términos accesibles a seres incapaces de imaginar cómo sería “ser nosotros”. (...) No quiere saber *para él* lo que es ser murciélago. Quiere saber *objetivamente* cómo es *subjetivamente* serlo. No sería suficiente para él haber tenido la experiencia de colocarse un “casco de bateador”, un casco con electrodos que estimule su cerebro hasta provocarle experiencias de murciélago y haber experimentado, por lo tanto, un “estado de murciélago”. Esto sería, en definitiva, tan sólo lo que sería para *Nagel* ser murciélago. ¿Qué lo podría satisfacer entonces? (Dennett, Daniel; Hofstadter, Douglas, 1983, pág. 529)

Si Nagel estuvo pensando en que la “fenomenología objetiva” podía satisfacerse *convirtiéndose* en el sujeto cuyo “*what is it like to be*” se desea conocer, entonces evidentemente ha pedido algo imposible. Quizás su especulación era más modesta; tal vez pretendía acuñar algún tipo de vocabulario fenomenológico que pudiera facilitar la comprensión intermodal entre personas que no poseen determinados sentidos. Ese vocabulario nacería de una reconceptualización de nuestras maneras de entender nuestras sensaciones. Sin embargo, él mismo pone en duda que una “fenomenología objetiva” pueda eliminar la brecha entre el fenómeno y el mundo, pues es posible que “algo quede excluido” (1983, pág. 520). En todo caso, una vez más, quizás le está pidiendo a las explicaciones y conceptos naturalistas más de lo que pueden dar. Pero, ¿qué es lo que pueden dar las explicaciones? ¿Qué es exigible en una explicación de la relación entre la fenomenología y los acaecimientos del mundo?

2.4. “Dennettizar” las explicaciones o “Nagelizarlas”.

Reductionism is not just a thesis about how the world is, it is a thesis about what the mind is like as well: namely, one that assumes that our various special-purpose representations can be regimented into the form of a single axiomatic system without loss of content or explanatory power. – Steven Horst, Beyond Reduction.

La diferencia de perspectiva entre Nagel y Hofstadter-Dennett no es sólo con respecto a la suficiencia o pertinencia explicativa del fenómeno. No es un desacuerdo acerca de cómo debe puentearse la brecha entre lo fenoménico y los acaecimientos del mundo; es un desacuerdo acerca del rol que debe cumplir una explicación para que cuente como explicación de un determinado fenómeno, y un desacuerdo acerca de qué es lo

que debe ser explicado. Para Daniel Dennett, *lo que debe ser explicado* debe ajustarse a una estrategia reduccionista propia de la ciencia natural y este es –el ámbito de la ciencia natural–, en rigor, el único dominio explicativo relevante. De hecho, una de las tareas de la investigación científica consiste en la búsqueda de una unidad explicativa de los fenómenos del mundo, y sería inconsistente con la práctica del científico natural afirmar que las propiedades de un dominio natural son inconmensurables o incompatibles con las propiedades de otro dominio natural. En estos dos aspectos (la reducción de los fenómenos a sus bases no fenoménicas y la unidad explicativa de todos los hechos del mundo) su propuesta no es esencialmente diferente de la que Locke acerca de los hechos del mundo natural. Para Nagel, en cambio, debe haber una explicación fenomenológica por derecho propio: lo fenoménico es un dominio específico con características irreducibles (o no completamente reducibles) al ámbito de la ciencia natural. En este caso, también encontramos a Locke: lo mental en sí mismo posee características propias, irreducibles (para Locke, dependían de la arbitrariedad divina). El desacuerdo está dado en términos de lo que cuenta como *explicaciones conceptualmente adecuadas* (CAE) del ámbito fenoménico. Horst (2007) define a una CAE como:

An explanation of A in terms of B is conceptually adequate just in case the conceptual content of B is sufficiently rich to generate that of A without the addition of anything fundamentally new (2007, pág. 33)

Si no se requieren de elementos fundamentalmente nuevos para explicar la naturaleza de los estados fenoménicos, entonces la estrategia reductiva podría proporcionarnos una CAE. En cambio, si lo fenoménico exige algún tipo de explicación *sui generis*, no conectada con hechos físicos básicos, entonces las CAE deberán buscarse en el propio ámbito fenoménico. Existen, sin embargo, tres tipos de CAE, cada uno más débil que el anterior:

The stronger form, which we may call a *Pure CAE* is found in syllogisms and mathematical constructions. (...)

A slightly weaker form of explanation is obtained when we augment the vocabulary of the explaining system with additional, but purely formal tools such as Descartes's application of algebraic tools to Euclidean geometry to create analytic geometry. (...)

There is at least one additional, and still weaker, form of CAE. Consider the explanation of how a given token system counts as a member of a functional kind: how a bit of

tissue counts as a heart, or how an arrangement of circuits counts as an AND-gate. In this case, the nature of the *kind* (“heart”, “AND-gate”) is not explained in the process. The anatomy of the heart does not explain *what is to be a heart*. However, if the functional kind is suitable well-understood, it becomes epistemically transparent how a system having *this set of physical* properties would thereby have *that set of functional* properties as well. (2007, págs. 33-4)

Si existe algún tipo de CAE en lo que respecta a la conexión entre el ámbito fenoménico y los hechos del mundo, será del tipo más débil: lo que una explicación puede hacer es mostrar de modo transparente la conexión entre propiedades físicas (del nivel n) y propiedades funcionales (del nivel $n+1$), aunque no podrá explicar la naturaleza de la clase (del nivel $n+1$) tal como es en sí misma. El corazón (como unidad funcional) debe “desaparecer” en el explanans si se lo explica mediante sus propiedades de bajo nivel inmediato. Esto es lo esperable en una explicación del ámbito fenoménico: si lo fenoménico se identifica con determinadas unidades funcionales (presuntamente localizadas en el cerebro), una CAE debe ser capaz de conectar estas unidades funcionales con propiedades físicas. El problema inicial es el de aceptar que los aspectos fenoménicos pueden identificarse con unidades funcionales, *sin pérdida de transparencia explicativa*. Es aquí donde Dennett y Nagel se separan.

Daniel Dennett ha popularizado la acuñación del verbo “quinear” (Dennett, Quining Qualia, 1988). Se quinea cuando se niega la existencia o la importancia de algo:

Quine, v: to Deny resolutely the existence or the importance of something real or significant” (Dennett, The Philosophical Lexicon, 1987)

Podemos apropiarnos de este método para acuñar verbos a partir del nombre de un filósofo:

Dennettizar (De Daniel Dennett, filósofo): Establecer que sólo puede darse explicación conceptualmente adecuada (CAE) de lo mental si se lo reduce al ámbito de las ciencias naturales.

La dennettización es una forma de reduccionismo que busca eliminar la fenomenología del ámbito explicativo. Para ello, debe negar dicha fenomenología para identificarla con estructuras funcionales y, con ello, debe rechazar la posibilidad de hacer

afirmaciones con sentido acerca de los propios fenómenos. La reducibilidad es un criterio de legitimación ontológica. (Horst, 2007, pág. 80)

Por el contrario, se la **nageliza** (de Ernst Nagel, filósofo) cuando se pide que el hecho sea explicado en los términos de la descripción del *explanandum*. Concretamente, cuando se pide que se explique el fenómeno *qua* fenómeno, dado que se rechaza que para tal fenómeno sea exigible *sin más* una explicación del tipo CAE en ciencias naturales: una importante fuente de CAEs *debe provenir del propio fenómeno qua fenómeno*.

Nótese que el nagelizador no tiene por qué negar que los fenómenos sean explicados por medio de una reducción a un nivel menor: simplemente dice que esa no es toda la explicación que se puede dar. Aun cuando los hechos fueran reducibles, esa reducción no deja fuera de juego al hecho reducido en tanto necesitado de algún tipo de explicación adicional de alto nivel.

Es importante resaltar qué debemos tener en cuenta por “explicación” y por “reducción” en estos contextos. Que un hecho H sea *explicado* significa que se ha encontrado otro hecho u otro conjunto de hechos C conectados con el hecho H de tal manera que quede explícita una *relación de dependencia necesaria* entre ellos, en el sentido de que la configuración de C es condición necesaria para H. Una explicación, además, es conceptualmente adecuada (CAE) si C es condición *necesaria y suficiente* para H. En este punto, tanto el dennettizador (reduccionista) como el nagelizador podrían estar de acuerdo. El desacuerdo sobrevendría acerca de cuáles son las condiciones necesarias y suficientes para explicar el hecho H (en este caso, la conciencia).

La explicación reductivista (dennettizadora) postula que C es el conjunto de componentes del sistema funcional que sirve de base a H, y que las propiedades y relaciones entre estos componentes son suficientes para la caracterización de H. En otras palabras: la explicación debe ser “de la parte al todo” (el “todo” de la conciencia se explica a partir de sus componentes basales: por ejemplo, las propiedades y

relaciones del soporte funcional nervioso), y se arguye que esta explicación es “sin resto”⁵².

Una explicación no reductivista, si bien puede aceptar que los componentes de un sistema son necesarios para explicar el fenómeno, no acepta que tales componentes sean suficientes. El caso de Locke es paradigmático: la arbitrariedad divina es un componente fundamental que no forma parte de la estructura basal, corpuscular, del universo. (En todo caso, en Locke, podría pensarse que la estructura corpuscular más la arbitrariedad divina conforman el conjunto C necesario y suficiente. Sin embargo, es importante destacar que ese conjunto está conformado por elementos basales heterogéneos y desconectado entre sí, y en rigor es difícil determinar en qué sentido la decisión divina es un elemento de la estructura basal de la realidad).

Otro tipo de explicación no reductivista postula que los estados mentales son supervenientes a los estados físicos, aunque la dependencia establecida entre ellos no es causal, sino de una *correlación necesaria*. No queda claro, de todos modos, si las explicaciones mediante estados supervenientes son en rigor explicaciones o si solo se limitan a encontrar correlaciones fiables entre ambos dominios. En palabras de Kim (2000, pág. 14): “We must conclude then that mind-body supervenience itself is not an *explanatory theory*; it merely states a pattern of property *covariation* between the mental and the physical and points to the existence of a dependency relation between the two. Yet supervenience is silent on the nature of the dependence relation that might explain why the mental supervenes on the physical”.

El dennettizador asume que la identificación de las CAEs acerca de lo mental con CAEs de la ciencia natural es un procedimiento normal en la historia de la ciencia: un fenómeno deja de ser explicado *qua* fenómeno y pasa a ser explicado en los términos de las CAEs de la ciencia natural. Por ejemplo, el fenómeno intuitivo e inmediato de la inmovilidad de la Tierra con respecto a los astros puede explicarse *sin resto* mediante modelos que sitúan al fenómeno desde una nueva perspectiva, desde la cual el parecer₁ inicial (“La Tierra está inmóvil y el Sol gira a su alrededor”) no se modifica, pero sí se modifica el parecer₂ (“Aunque no siento el movimiento de la Tierra y tengo

⁵² “[T]he characteristic features of broadly reductive explanations are that they are: I. Part-whole explanations (...) II. Explanations without remainder”. Horst, 2007, pág.31.

la impresión de que es el Sol el que se mueve [parecer₁], de acuerdo con el conocimiento acerca de leyes gravitacionales y de la dinámica celestial, me parece que la Tierra es la que se mueve” [parecer₂]). La estrategia del dennettizador consiste en encontrar un conjunto de CAEs con suficiente poder explicativo como para que cada fenómeno del tipo parecer₁ pueda ser explicado sin resto por un conjunto de proposiciones de parecer₂. Es importante, de todos modos, enfatizar que el parecer₁ no se elimina aunque haya una CAE de ciencia natural que explique sin resto el fenómeno.

Con respecto al status de la experiencia consciente, el dennettizador asume que la conciencia dejará de ser un misterio cuando se logren avances conceptuales y científicos análogos a los ocurridos durante la disputa geocentrismo versus heliocentrismo, vitalismo versus mecanicismo y creacionismo versus evolucionismo. Estas analogías no son simples ejemplos: son casos paradigmáticos de éxito científico en los cuales un fenómeno o un elemento teórico que se consideraba fundamental e irreducible pudo ser explicado sin resto por medio de CAEs propias de la ciencia natural.

El nagelizador puede replicar que tal reducción es operable si se la realiza sobre elementos teóricos inciertos (el creacionismo, lo mismo que el flogisto, no implican ningún tipo de *fenómeno positivo*; ambos conceptos son postulados teóricos, que se someten a una reducción cuando se problematiza la efectiva existencia de tal postulado) o sobre fenómenos que son producto de la perspectiva de una conciencia (la apariencia de que la Tierra no se mueve). Lo que no pueden es reducir a la propia conciencia como elemento originario de todos los fenómenos. Es posible reducir los fenómenos –que son fenómenos *para* una mente- a CAEs propias de la ciencia natural, en las que el fenómeno *qua* fenómeno desaparece. Pero es por lo menos sorprendente que la propia conciencia, como fuente originaria de fenómenos, sea explicada con prescindencia de la característica más saliente: su capacidad de representación fenoménica. Quizás las explicaciones por medio de CAE de ciencias naturales deban eliminar los fenómenos en tanto fenómenos para la mente, pero no pueden eliminar a la mente como fenoménica en sí misma. En otras palabras, pueden explicar que a una mente se le aparezcan fenómenos y pueden mostrar que esos fenómenos son

necesaria y suficientemente causados por sucesos de bajo nivel. Pero no pueden mostrar que la propia mente es un fenómeno superveniente, porque la demostración de este hecho sería una demostración de la mente qua-no-mente. Mostrar eso es generar la experiencia de que lo que parece ser esto que estoy siendo en realidad no es lo que parece. Esto último, desde luego, no es un problema para el denetizador: ¡De hecho, la mente no es lo que parece!

Decir que algo es fenómeno equivale a decir que es “dado a una mente”. Como Frege notaba (v. § 1.8), no puede haber representación si no hay portador. Todo fenómeno es un fenómeno mental y, por ello, un fenómeno como la experiencia de que la Tierra está inmóvil no sólo habla acerca del hecho natural de que no podemos percibir el movimiento de nuestro planeta, sino que también habla acerca de que estamos hechos de manera tal que la experiencia de habitar un planeta en movimiento nos es dada *como si* el planeta estuviera inmóvil. El fenómeno habla, fundamentalmente, de la *perspectiva* en que está situado un ser consciente, tanto en lo que respecta a su situación espaciotemporal (su entorno) como a sus estructuras corporales, perceptivas y cerebrales. Una CAE adecuada para explicar los fenómenos debe eliminar la perspectiva de primera persona, para reemplazarla por una perspectiva neutral, en la que no se juzga acerca del modo en que se presenta el hecho explicado. La neutralidad con respecto a la perspectiva prescinde de la noción de portador o, en el caso de que necesite incorporarlo, lo hará sin referencia a sus estados fenoménicos en sus *explanantes* (los estados fenoménicos son siempre parte del *explanandum*).

En la introducción de esta segunda parte distinguimos dos nociones de fenómeno: una, la que se refiere a las características globales de la conciencia y otra, la referida a características locales o puntuales. Las características globales nos remiten a una pregunta fenoménica radical: “¿Cómo es ser tal o cual individuo de tal o cual especie?”. Por supuesto, esta ya famosa pregunta de Nagel que dio origen a toda una tradición de preguntas del tipo “what is it to be like”, no es la única posible en lo que respecta a las características globales. Una pregunta quizás más radical nace de la propia experiencia de ser uno mismo: ¿por qué yo soy yo y estoy atado a este cuerpo? ¿Por qué cada mañana soy *este*, y no más bien una hormiga, o un elefante, o una piedra? ¿Qué hace que a este cuerpo le acompañe esta conciencia y no otra, o mejor

ninguna?⁵³ Hay un “sentirse como yo”: el tinte de las propias experiencias, la historia individual que acompaña al cuerpo, el modo en que se presentan los fenómenos, la reacción ante ellos. Hay miles de millones de individuos conscientes en la Tierra: ¿por qué yo soy *solo* este, *continuamente* este y *justo* este y nunca otro? Estas son preguntas para las que no parece haber una respuesta clara desde la estrategia dennettizadora.

Por el contrario, las características locales parece que pueden explicarse sin problemas mediante CAEs de ciencia natural. “¿Por qué veo un anillo rosado en la contraportada del libro?”, “¿Por qué veo rojo y no verde?”, “¿Por qué cuando una superficie tiene una temperatura superior a los cuarenta grados centígrados se percibe la sensación de calor?”. Reduciendo los estados cualitativos de la conciencia a sus estados funcionales relevantes, (o, para utilizar una terminología moderna, reduciendo cualidades secundarias a cualidades primarias) parece posible obtener una explicación sin resto del estado fenoménico. El estado fenoménico se reconceptualiza como un conjunto de acaecimientos de un determinado tipo.

Por supuesto, es crucial para nuestra supervivencia cotidiana suponer que detrás de ciertos tipos de fenómenos hay regularmente ciertos tipos de acaecimientos. Cuando decimos “El auto rojo que cruza la calle es idéntico al auto rojo que cruzó hace diez minutos” no pretendemos comparar fenómenos, sino acaecimientos. Cuando decimos “La melodía que pensé hace diez minutos es idéntica a la que estoy teniendo en este momento”, pretendemos comparar fenómenos. La explicación dennetizada asume que en los dos casos existen acaecimientos subyacentes, y la comparación debe realizarse, en rigor, entre tales acaecimientos. Por lo tanto, en el caso de la comparación de melodías, o bien a) se debe reducir el componente fenoménico para que la comparación se efectúe entre acaecimientos observables en tercera persona; o bien b) no hay actividad en absoluto; tan solo la *apariencia* de un suceso mental. La respuesta

⁵³ D. R. Hofstadter, ante la afirmación de que nuestro yo elige una ruta de identidad entre una multitud de ramificaciones del cosmos, en las cuales existen otros “yos” que han elegido ramificaciones del cosmos diferentes, se pregunta: “¿Por qué este sentimiento unitario de ser yo se propaga por *esta* rama cualquiera más bien que por otra? (...) ¿Por qué mi sentimiento de mí mismo no acompaña a mis otros yos cuando se separan para seguir otras rutas? ¿Qué liga el carácter de yo al punto de vista de este cuerpo que existe en esta rama del universo en este momento del tiempo? La cuestión es tan fundamental que llega a resistirse a una formulación clara en palabras.” (Morowitz, 1983, pág. 66)

b) ha sido dada por Dennett, como hemos visto más arriba, en § 2.3. Pero, ¿cómo es posible que uno tenga la *apariencia* de un suceso mental? Que algo le parezca a uno es, precisamente, un suceso mental. Por lo tanto, ni siquiera debería poder decir legítimamente que me *parece* que estoy viendo un anillo rosa. “Parece haber fenomenología”, dice Otto. Dennett declara que de ese parecer no se deduce que haya fenomenología. Pero hay un parecer, y este parecer debería explicarse mediante una “fenomenología objetiva”, dice Nagel. ¿Qué ocurre con este parecer? ¿Es este parecer *dennettizable*?

2.5. Los *qualia* no son *lo que parece*. (O: ¿le parece a usted que a mí me parece que el césped es verde?)

Una de las palabras más popularizadas para nombrar los estados fenoménicos es la palabra ‘quale’ (plural: ‘qualia’). Los qualia consisten en las cualidades subjetivas de la experiencia. El color rojo borgoña de la mesa de la cafetería produce una impresión subjetiva, difícilmente comunicable, notoriamente privada y directamente aprehensible para una conciencia. Usted no necesita hacer una inferencia para saber que está teniendo el quale de rojo borgoña, o para conocer que le duele el pie derecho. Los qualia tienen las cuatro características de las vivencias enumeradas en § 1.1: privacidad, transparencia, irreducibilidad e incorregibilidad (García Carpintero, 1996, pág. 73). Los estados cualitativos se conocen de un único modo: vivenciándolos. Son el contenido fenoménico de cada vivencia subjetiva, y los estados conscientes se caracterizan por estas vivencias.

En *Quining Qualia* (1988) Daniel Dennett declara: “Qualia is an unfamiliar term for something that could not be more familiar to each of us: the *way things seem to us*”. Así definido, el quale es idéntico al *modo subjetivo como nos parecen las cosas*. Pero, ¿es legítimo establecer esta identidad entre *quale* y *parecer*₁?

El rojo borgoña de la mesa que estoy viendo en este momento es un tipo de estado cualitativo de mi conciencia, y en ese sentido –en tanto es un contenido de conciencia– es interno. Sin embargo, *el modo en que se me aparece este rojo* no es como algo

interno, sino más bien como si estuviera *en la mesa*. Al menos en lo que se refiere al aspecto interno, el quale no es lo que parece, pues no parece₁ como si fuera algo “dentro” de la propia mente. John Gibbons (Gibbons, 2005) analiza esta peculiar propiedad de los qualia, de “estar en la mente”, a diferencia del “parecer”, que proyecta esos estados cualitativos no hacia una interioridad subjetiva, sino hacia un afuera:

We are looking for reason to believe in qualia, or the redness that is in the mind. Let us start with introspection. First, we need to distinguish two questions. The first question is this: Where is the closest red thing to you? I only raise this question to distinguish it from the second question and then set it aside. Here is the question that matters: Where does the closest red thing seems to be? If you are not visible bleeding, and you are like the rest of us, it will seem to you as though the closest red thing is at some distance from you. It seems to be outside of you. It does not seem to be inside of you. Of course, it does not follow from this that it really is outside of you. (...)
Where is the redness seems to be? “There!” you say, pointing to the external world. Of course, it might not really be there. That is just where it seems to be. (2005, pág. 413)

Si se nos pregunta, sin intención filosófica, dónde vemos el color rojo, señalaremos en primer lugar el *espacio exterior*, la superficie donde parece estar el color. Las sensaciones poseen un componente intencional; señalan algún espacio exterior o alguna parte del cuerpo. La sensación más el componente intencional conforman el *parecer₁* de la sensación. Supongamos una sensación ficticia: un dolor de cabeza sumamente intenso, pero que parece localizarse no en la propia cabeza, sino *más arriba* de ella (unos veinte centímetros, o, si se desea exagerar, cien kilómetros). Podríamos llamarlo *metacraña* (En lugar de migraña). En este caso hay un desacople entre el componente puramente cualitativo del dolor y el componente intencional. El individuo afectado de metacraña tiene un dolor real, pero ese dolor real tiene una *intencionalidad fantasma*⁵⁴: al individuo le parece₁ que su dolor no está en la cabeza, sino en un órgano que no posee (que nunca poseyó) ubicado muy por encima de su sombrero. El *parecer₁*, entonces, está conformado por ese dolor inmediatamente percibido (el quale) más la inusual ubicación espacial de lo dolorido. Pero, precisamente, la “localización fantasma” de este dolor parece contradecir la propia

⁵⁴ El llamado “dolor fantasma”, ¿es realmente un dolor fantasmal, o es un dolor *real* con una intencionalidad fantasma?

definición de quale, pues el dolor no está localizado “en mi mente”, sino en alguna parte por encima de mi cabeza.

Sin embargo, a pesar de la afirmación de Gibbons, no todos los estados cualitativos tienen un parecer dirigido “hacia afuera”. El color rojo “más cercano” que se pudiera señalar podía ser –en lugar del rojo de la mesa- un fosfeno. Si cerramos los ojos y nos los frotamos, veremos una danza de colores. En ese caso, se podría decir que el rojo que tenemos más cerca es ese que vemos con los ojos cerrados. Es difícil señalar una localización espacial para los fosfenos. Diremos algo así como que “parece que hay un rojo danzando ante mis ojos”. (Más adelante trataremos específicamente este caso, v. Block, [2003]) Quizás, exceptuando estos fenómenos marginales, debemos aceptar que el parecer más común de las cosas rojas es un parecer externo. Lo mismo vale para el dolor. No es común escuchar que alguien sienta un dolor “en la mente”, a menos que lo diga de modo metafórico. No solemos señalar nuestra cabeza cuando señalamos lo que estamos viendo o sintiendo, tal como lo ilustra Wittgenstein en § 275:

Mira el azul del cielo y dite a ti mismo «¡Qué azul es el cielo!»- cuando lo haces espontáneamente- no con intenciones filosóficas no te viene la idea de que esa impresión de color te pertenezca sólo *a ti*. Y no tienes reparo en dirigirle esa exclamación a otra persona. Y si señalas algo al tiempo que dices las palabras, se trata del cielo. Quiero decir: no tienes el sentimiento de señalar-a-ti-mismo que acompaña al ‘nombrar la sensación’ cuando se medita sobre el ‘lenguaje privado’. Tampoco piensas que propiamente deberías señalar el color no con la mano, sino sólo con la atención. (Considera lo que quiere decir «señalar algo con la atención») (1988, pág. 235)

Nuestro parecer₁ acerca de la mayoría de nuestros fenómenos puede ser señalado hacia afuera y no “hacia adentro” de la propia conciencia. Presuntamente, el quale es algo que está “adentro”, pero el parecer₁ nos muestra como si ese estado cualitativo de nuestra conciencia no estuviera en el interior de ella, sino “afuera”.

Gibbons utilizará tres argumentos para separar la noción de “quale” de la noción de “parecer”: el argumento de la *inversión de espectros*, el de la *introspección* y el de la *ilusión*. En los tres casos, el aspecto cualitativo se puede entender como separado de lo que aquí hemos llamado parecer₁. Gibbons mostrará que los *Ways things seem* no son,

como dice Dennett en Quining Qualia, sinónimos de *quale*. Sus argumentos partirán de intuiciones básicas de la experiencia.

El primer argumento es el de los *pareceres* invertidos. Si aceptamos que qualia y *parecer* (*ways things seem*) son diferentes, los pareceres invertidos son imposibles. Partirá de una primera hipótesis:

(Hype) As far as their colors are concerned, fire hydrants seem to Laverne the way grass seems to Shirley (2005, pág. 398)

Los carros hidrantes son típicamente rojos y el césped es típicamente verde. Pero eso no es lo importante. No se está afirmando que a Laverne le parezcan verdes los carros hidrantes, o que a Shirley le parezca rojo el césped. Simplemente tienen *el mismo parecer₁* acerca de hechos que poseen estados cualitativos invertidos, sea cual fuere el criterio de identidad de ese parecer. Podemos mantenernos neutrales con respecto a la naturaleza de ese estado cualitativo; nos centramos en el *parecer₁*. Asumamos que Shirley y Laverne no son indistinguibles física, intrínseca o funcionalmente, pero sí lo son con respecto a este específico *parecer₁* (2005, pág. 399).

Hype es imposible por autocontradictoria.

En primer lugar, ellas utilizan los términos “rojo” y “verde” como nombres para los mismos colores. Ambas dirán que “los carros hidrantes son rojos”, aun cuando el *parecer₁* de ambas fuera distinto. Ambas tienen las mismas creencias acerca de los carros hidrantes y ninguna de las dos está en mejor posición para conocer el color de los carros hidrantes. Supongamos, además, que no mienten y que no se equivocan con respecto a lo que les parece₁ (2005, pág. 401). Es necesario suponer que uno puede conocer lo que ocurre en su propia mente. Ambas creen que los carros hidrantes son rojos y que el césped es verde: en lo que respecta a los carros hidrantes, a ambas les parece₁ que son del mismo color (rojo). En lo que respecta al césped, ambas coinciden en que les parece₁ que es del mismo color (verde), el cual es distinto al de los carros hidrantes. De aquí se sigue que a Laverne no pueden parecerle₁ los carros hidrantes del modo en que le parece₁ a Shirley el césped, (2005, pág. 402) lo cual es la negación de Hype. Tomada literalmente, Hype desemboca en su propia negación.

Hype presupone que los colores se mantienen fijos, mientras que los pareceres₁ (*ways things seems*) pueden diferir, como si hubiera un específico estado cualitativo (un *quale*) que se mantiene mientras cambia el parecer₁ relativo a ese estado cualitativo. Pero, ¿por qué es necesario aceptar esto?

Supongamos otra posibilidad: si somos llevados a la Tierra invertida de Block (Block, *Inverted Earth*, 1990), en la que todos los hechos son idénticos a los de nuestra Tierra, con la excepción de que todos los colores están invertidos. Supongamos que se toman medidas para contrarrestar esta inversión: se nos colocan lentes que invierten los colores, de modo que vemos el color que veíamos en la Tierra original. (Gibbons, 2005, pág. 407) Cuando vemos el cielo y decimos “el cielo es azul”, estamos hablando de ese cielo que estamos viendo en la Otra Tierra, pero la referencia de “Azul” está fijada de acuerdo a cómo aprendimos a usar el término en la Tierra. “Azul” puede describirse, típicamente, como “el color del cielo”. A medida que pasa el tiempo, vamos a cambiar subrepticamente la referencia de nuestros términos de estados cualitativos. “Azul” pasará a designar “El color del cielo en la Otra Tierra”. Algo similar puede ocurrir con la referencia de descripciones como “Mi mejor amigo”. Si nuestro vecino José⁵⁵ cumple con esta descripción, en la Otra Tierra me encontraré con Otro José y, erróneamente, le adscribiré esa descripción. Sin embargo, con el tiempo puede que el Otro José se convierta en mi mejor amigo. En ese caso, la descripción que inicialmente era errónea (y no lo sabíamos) pasó a ser correcta.

En la Otra Tierra hay carros hidrantes, y decimos de ellos que son rojos. No nos damos cuenta de que tenemos puestos los lentes inversores, de modo que decimos *lo que nos parece*. “Cualquiera sea el color de los carros hidrantes, a mí me parecen₁ rojos”, decimos. Ser transportados a otro planeta no obstruye nuestro acceso privilegiado a los hechos de cómo nos parecen las cosas (2005, págs. 408 - 9). Sin embargo, habremos cometido un error. En la Otra Tierra, “rojo” significa lo que para mí, anteriormente, en la Tierra, significaba “verde”. De modo que si digo que los carros hidrantes y los tomates maduros *me parecen₁* rojos, estoy diciendo (sin saberlo) que

⁵⁵ Gibbons utiliza como ejemplo a “Laverne”. Preferimos cambiar el nombre para no confundirlo con el de la hipótesis (Hype) dada anteriormente.

me parecen₁ verdes. (2005, pág. 409) Mi parecer₁ ha cambiado sin que me haya dado cuenta.

Parece que el parecer no puede invertirse, y puede cambiarse sin que nos demos cuenta. ¿Son estas las características de los qualia?:

If qualia are ways things seem, then qualia are not what they seem. They do not supervene on neurophysiology; they can change without your noticing, and inverted qualia are impossible. If qualia are not ways things seem, then you should be very worried about what justification you can have for believing that qualia, unlike ways things seem, are intrinsic, invertible or what have you. (2005, pág. 410)

¿Debemos creer en la existencia de los qualia, como algo diferente de los pareceres₁ (*ways thing seem*)?

El segundo argumento apelará a la introspección (2005, pág. 412). Como dijimos al principio de esta sección, si señalamos nuestro *quale* mediante introspección, diremos que algo rojo nos *parece* estar a cierta distancia, aunque el quale rojo es algo interno, intrínseco a nuestra experiencia. De ese modo, la introspección nos revela a un quale como interno, pero también a un parecer de que ese quale no está ubicado en un espacio interno de representaciones, tal como ilustrábamos con la cita de Wittgenstein, § 275: “no tienes el sentimiento de señalar-a-ti-mismo que acompaña la sensación”. Si sostenemos, con todo, que hay algo como “la rojedad en la mente”, dicha “rojedad” no puede basarse en *lo que nos parece₁*. “It visually seems to you that the redness is in the world and not in the mind. This is some evidence, not conclusive evidence, but some evidence that there is no redness in the mind” (2005, pág. 415)

Los qualia del rojo están en la mente; son intrínsecamente internos. El parecer coloca al rojo afuera. Una vez más, parecer y quale están desacoplados.

El tercer argumento es el de la ilusión. (2005, pág. 420) Cuando más arriba vimos los ejemplos de las ilusiones ópticas, hablamos de la ilusión de un anillo rosa, de un círculo verde y de una cuadrícula curvada. No parece haber un obstáculo para imputarle colores o, en todo caso, “ideas” derivadas de cualidades secundarias a “la mente”. Pero es difícil admitir que hay algo así como la “anillidad” o la “cuadradez” en la mente. Dicho en otras palabras, parece mucho más fácil admitir que, cuando se alucina

con un color, el color sea un hecho interno. Pero cuando se alucina con un objeto material, sólido, tridimensional –es decir, con un objeto “hecho” a partir de cualidades primarias-, no es fácil aceptar que esas características puedan ser hechos internos. Supongamos que usted alucina con una roca roja:

If you are hallucinating a red rock, neither the redness nor the rockness is in the world. So they must be in the mind. By now, I will assume, you no longer think you have rocks in your head. But then, you are faced with the difficult question of why the argument for mental redness works, if it does work, while the argument for mental rockness does not. There must be a distinction between properties like redness and properties like rockness(2005, pág. 420)

Es difícil aceptar propiedades como la “roquedad mental”, el “carácter de ser una roca” puramente mental. Sin embargo, parece más fácil aceptar la “rojedad-en-la-mente”. Ahora bien, una cualidad secundaria no se presenta aislada: tener la alucinación de rojo implica un espacio donde ese rojo se da; debe tener una forma u otra: “You just can not have an exemplification of redness, any kind of redness, without its having some kind of shape” (2005, pág. 421) Si la rojedad tiene una forma, debe estar en un espacio. Si no es en el espacio físico, debe ser en un espacio mental.

Sin embargo la admisión de un espacio mental para la rojedad nos lleva a una admisión sumamente extraña e indeseable:

When you hallucinating a red rock, the rock, the redness, and the roundness are not really out there in the world. And this apparent out-thereness, or apparent objectivity is as much a feature of your experience as the apparent redness and roundness. If the apparent redness is a kind of redness, subjective redness, then maybe the apparent objectivity is a kind of objectivity, subjective objectivity. No matter what else you believe, no one should believe in subjective objectivity. (2005, pág. 421)

Una roca alucinada parece proyectada en un espacio de objetividad, una *objetividad aparente*. Nos parece que la roca existe, como la daga de Macbeth. Pero esa objetividad es falsa; nos *parece* que es así, pero no es. Como el anillo rosa, parece *objetivamente* proyectado en la contraportada del libro. Cuando Otto preguntaba “¿Dónde está?” el anillo rosa, quizás estaba azorado por la persistente apariencia de objetividad de la imagen que veía. *Quizás preguntaba por el espacio donde residía ese anillo*, un espacio proyectivo puramente ilusorio. Dentro de la ilusión, está la *sensación*

de objetividad de ese espacio. La noción de “objetividad subjetiva” suena contradictoria, y no parece muy alejada de las nociones de “aparición real”, “fenomenología real” de Dennett y la de “fenomenología objetiva” que proponía Nagel.

La consecuencia de este tercer argumento es que no tenemos un principio claro para distinguir a los qualia de otras propiedades fenoménicas como la “roquedad”. Si *qualia* es un término latino para referirse a los estados *cualitativos*, ¿no podríamos utilizar el término *quidia*, para referirnos a un estado *de sustancia* (relacionada con cualidades primarias) de los fenómenos mentales?⁵⁶. El quale alucinado debe ser mental (interno), pero ese quale *parece*₁ proyectado en una “objetividad subjetiva” (externa). Una vez más, qualia y *what it seems* están desacoplados.

Ninguno de los tres argumentos sustenta, además, la afirmación de que existen los qualia. No se niega concluyentemente, pero hay una sugerencia de que el carácter *interno* de los estados cualitativos de conciencia es sospechoso. Tanto en los casos de una experiencia verídica, como de una experiencia ilusoria, los caracteres de la experiencia parecen proyectados en un espacio externo, y el contenido de nuestro *parecer*₁ se identifica (o *parece*₁ identificarse) con ciertas propiedades externas. Los qualia se vuelven una ociosa criatura de existencia puramente privada e interna, que sólo surge después de un análisis filosófico, pero que no aparece nunca en el *parecer* inmediato ordinario. ¿Necesitamos de los qualia, esto es; es necesario postular el carácter privado, interno e inefable de los estados cualitativos de conciencia, después de todo? Si encontramos argumentos para eliminar a los *qualia* en tanto que componentes fenomenológicos internos, entonces el camino para dennettizar las explicaciones acerca de tales estados cualitativos parece estar abierto, y la reducción de los estados fenoménicos a estados funcionales / mecánicos podría operarse sin problemas, pues ya no se postularía un conjunto de insondables entidades fenomenológicas internas, sino pareceres₁ inmediatamente proyectados hacia el mundo. Un problema que debemos enfrentar es: si conseguimos la eliminación de los qualia, esto es, de los estados internos, ¿conseguiremos, a su vez, eliminar la

⁵⁶ Desde luego, este es un juego especulativo, pero, si aceptamos los qualia, ¿por qué no aceptar *quidia*, *quanta*, *quisia* (de *quis* = quién) -o cualquier otro término derivado de un pronombre interrogativo latino- mentales?

fenomenología? Después de todo, el parecer₁, proyectado hacia afuera, es fenomenología, cuya sede no parece ser una misteriosa interioridad, sino un “espacio de objetividad subjetiva”, proyectado hacia un ámbito externo que difiere de la propia interioridad de la mente (el rojo de la mesa de la cafetería no es un misterioso e inefable quale “dentro de mi mente”, sino algo que parece dado en la propia mesa). Sin embargo, es posible que *ya* poseamos explicaciones que reducen exitosa y completamente los estados internos, fenoménicos, a estados externos, funcionales. Esa es, al menos, la afirmación de Jeff Foss que analizaremos en el próximo apartado.

2.6. Dennettizando la explicación sobre los qualia: la explicación de Jeff Foss.

El dennettizador asume que todo lo que se requiere para explicar el fenómeno es una explicación de los procesos funcionales de bajo nivel subyacentes. Por “bajo nivel” se ha de entender una explicación que descomponga al fenómeno en un nivel funcional *inmediatamente* menor. No se habla de una explicación por medio del más bajo nivel posible, sino de uno que, dada su inmediatez con el nivel de la experiencia consciente, sirva como puente nomológico inmediato para inducir que *los fenómenos*, qua fenómenos, *sólo* son el producto de la acción de esas estructuras funcionales.

Jeff Foss, (Foss, 2009), nos presenta un aspecto claramente *dennettizador* del fenómeno de los qualia:

It is commonly maintained that qualia (in their essential nature) cannot be explained by science even *in principle*. Against this thesis I will argue that *in fact* we already have perfectly normal scientific explanation of qualia. To be specific, I will focus on the most difficult case, the case of colour qualia (2009, pág. 479)

La explicación *ya está dada*, de modo que no tiene sentido hacer preguntas adicionales o proponer “sofisticados argumentos” propios de los qualófilos. Foss remarca que lo que se entiende por “explicación de los qualia de color” tiene diferentes significados para diferentes personas (2009, pág. 480), de modo que su mira está en *explicar el modo en que se ven los colores*, es decir, el *parecer*₁. Su argumentación apoya algunas de las sugerencias de Gibbons:

John Gibbons (2005) presents intriguing arguments to the conclusion that qualia cannot be identified with the way things seem (pp. 397-411) (...) To the extent that qualia are defined in terms of the way things seem or look (or even what it is like to experience them), Gibbon's argument seems to leave them undefined, and hence to encourage skepticism about them. His explicit conclusion, however, is more modest: "I have not shown that there are no qualia... I have shown that it is not obvious that there are qualia" (425) (2009, pág. 506n)

Recordemos que en el apartado anterior habíamos encontrado argumentos para separar la noción de 'quale', el estado fenomenológico interno de la conciencia, de la noción de 'parecer₁'. El nagelizador mantendrá que una explicación acerca de los mecanismos de la conciencia debe explicar el contenido fenomenológico, esto es, *qualia* o *pareceres₁* en tanto que componentes fenomenológicos irreductibles. Foss dará una detallada descripción de *cómo se ve* (parece₁) el amarillo.

Y1) Yellow is the *brightest* of the colours. Blue and green and the rest seem intrinsically darker colours. Blue and green can be bright, as in the case of a bright blue sky, or bright blue light, but they purchase their brightness by sacrificing their hue. (...)

Y2) Yellow has specific, qualitative, *colour – relationships*. It is most similar to orange, and least similar to blue. Given the intervening colours there is a similarity continuum: orange, red, green, purple, blue.

Y3) Yellow is *atomic*. It is not an obvious compound in the way in which pink or purple are; the first a blend of red and white, the second a blend of red and blue.

Y4) Yellow has a specific *chemistry*: it can form compounds with some hues but not others. It combines with red to yield orange, or with green to yield chartreuse. But it cannot combine with blue. One cannot see, or imagine, a yellowish blue. (A juxtaposition of tiny blue dots and tiny yellow dots looks green, and mixing blue paint and yellow paint yields green paint. However, green does not *look* like a compound of yellow and blue. True compounds, like pink and purple, look like their components: pink looks reddish *and* white – ish, purple look bluish *and* reddish. But green does not look blue, and it does not look yellow. (2009, pág. 480-1)

Y1 – Y4 conforma el *explanandum*. Es un resumen detallado de *lo que nos parece₁* el amarillo o, más específicamente, de *lo que vemos* cuando vemos amarillo. El *explanans* general invoca a las *células cónicas* de la retina y el *proceso oponente* para la visión de los colores:

The human retina contains *cone cells*, which are sensitive not only to the overall brightness of visible light (as are the rod cells) but to its wavelength as well. There are three types of cone cell, each maximally sensitive to a specific wavelength, located at

440, 550 and 570 nanometres, with the corresponding cone types designates “blue”, “green” and “red”, (B, G, and R)

The output of the cones is then subjected pairwise to *opponent processing*. Whereby two inputs compete, with only the stronger transmitted as output. So where R and G cells provide input to an opponent process, the output is either some level of R activity, or G activity but not both. (...)

Cone output is routed into two separate opponent process, namely R versus G, and B versus the sum of R and G. The double opponent process constitutes a two dimensional space, with the $R \leftrightarrow G$ axis orthogonal to (independent of) the $B \leftrightarrow (R+G)$ axis. (2009, pág. 481)

El proceso de los pares oponentes puede compararse al de una palanca⁵⁷ (o un *joystick*) que tuviera cuatro posiciones básicas posibles (figura 9):

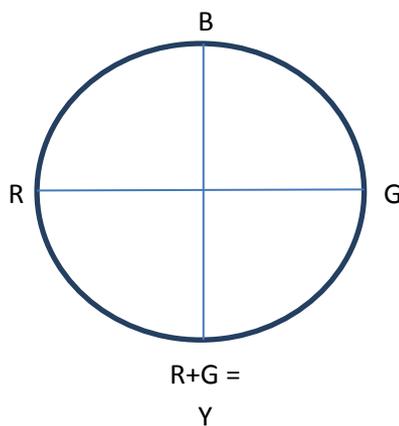


Figura 9. Esquema de los pares oponentes de los colores unitarios.

⁵⁷ Esta comparación no está presente en el texto de Foss; es una analogía ilustrativa que, a nuestro juicio, puede dar cuenta de manera intuitiva de las relaciones de similitud en la visión de los colores. A su vez, como se desarrollará más abajo, también es útil para explicar por qué ciertas combinaciones de colores *no* pueden darse.

Si la “palanca” se activa cerca del final del punto R, se ve el color rojo. Si lo hace cerca del final del punto G, se ve color verde; en el final del punto B se verá azul, y en el punto R+G, se verá amarillo. Para que pueda verse un color *puro*, es decir, uno que no sea producto de una mezcla reconocible de estos cuatro colores, el eje ortogonal debe estar en el punto medio, de modo que no interfiera agregando algún tipo de tonalidad. Si el eje ortogonal no está en el punto medio, tenemos colores mezclados: amarillo y rojo dan como resultado *naranja*; amarillo y verde da *verde lima* (chartreuse), azul y rojo para *púrpura*, azul y verde para *aguamarina*. (2009, pág. 481) En todos los casos de colores mezclados, podemos reconocer fácilmente sus componentes. Si imaginamos a este proceso como una “palanca” que tiene cuatro puntos, los colores mezclados deben pensarse como movimientos de dicha palanca a través del perímetro del círculo. Ese desplazamiento provocará la *mezcla* (Figura 9):

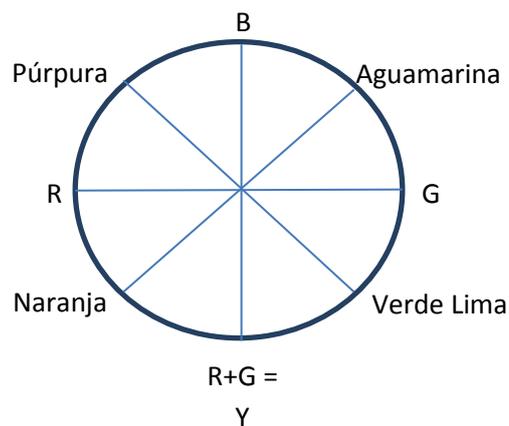


Figura 10. Esquema de pares oponentes de colores unitarios y binarios.

Si imaginamos que una palanca se puede mover en todo el continuo del perímetro del círculo, tendremos una cantidad infinita de puntos entre cada una de las secciones, lo que nos daría como resultado la visión de tonalidades diferentes, más o menos parecidas a los colores puros según la cercanía del punto con los ejes R-G o B-Y.

Una vez incorporados estos elementos como *explananda* generales, se pueden explicar específicamente los *explananda* Y1 – Y4:

This theory implies Y1-Y4 (here labelled, qua explananda, as E1-E4, and presented in reverse order) as follow:

E4) Yellow has a specific colour chemistry. This follows from the physiological processes of colour vision. Since yellow results from an opponent process with blue, it cannot blend with blue. Not only do we never see such a compound, we cannot even imagine it, because imagining a colour requires auto-stimulation for neural colour processes. (...)

E3) Yellow is atomic. This is because yellow is one of the four basic hues created by human colour processing. It is at one extreme of a single opponent process.

E2) Yellow has specific, qualitative colour relationships. It is least like blue, since it is opposed to blue. It is most similar to orange because orange is a compound of yellow itself (everything being self-similar). The nearly equal similarity of red and green with yellow is readily explained by the orthogonality of the red green axis with the yellow-blue axis: the independence of the red-green process from the yellow-blue process implies that red and green are equally similar (or dissimilar) to yellow. Many find red more similar to yellow than green. Red and yellow are said to be “warm” or “approaching”, while blue and green are “cool” or “receding”. (...) The lesser similarity of yellow with purple (than with either red or green) is a direct consequence of the fact that purple is a compound of blue, which is most dissimilar from yellow.

E1) Yellow is the brightest of the colours. This is because it results from an unfair opponent process in which the R and G cones gang up on the B cones. It is not surprising that the sum of R and G activity (yielding yellow) is bright, nor that the B activity (yielding blue) only triumphs when the light is low, when both R and G must be weak. So yellow is paradigmatically bright and blue is paradigmatically dark. A surprising consequence is that the brightness of yellow is, in a sense, an *illusion* born of the nervous system’s confusion of colour with brightness. There is such a thing as dark yellow, but we normally term this “brown”. (2009; págs. 481-2)

Estos cuatro puntos explican de manera suficiente *todo* lo relacionado con la visión del color amarillo. Si usted tiene la impresión de que su “sensación de amarillo” no está suficientemente explicada, entonces desea que esta explicación se *nagelice*: la “fenomenología del amarillo”, lo que es para usted el rico aparecer del amarillo de una margarita, necesita *algo más* que una explicación basada en términos tales como *células de la retina* y *pares oponentes*. El resto del artículo de Foss se dedica a explorar con elegancia, sobriedad y profundidad, trece posibles objeciones a esta explicación a la que calificamos como “dennettizada”, pues asume no sólo que los puntos E1-E4 conforman una CAE suficiente del *parecer*₁ de la visión del color amarillo, sino incluso que dicha explicación no presenta ningún tipo de desafío, no requiere la apelación a principios nuevos o fundamentales ni deja espacios inexplicados. No nos detendremos en todas las objeciones; sólo en aquellas en las que se cuestione específicamente la

naturaleza de la explicación. La objeción 2 podría ser propuesta por el nagelizador, quien tiene la sensación de que el explanans no ha explicado lo que debía:

Objection 2: The explanation transmogrifies (sic) the original explanandum, rich and sensuous qualia, into something more suited to its own devices, namely the skeletal colour-logic derived from the explanans itself. It amounts to a classic bait-and-switch manoeuvre: lured by the promise of an explanation of qualia, we get only an explanation of the scientific ersatz, “qualia”. (2009, pág. 484)

Foss responde⁵⁸:

- En primer lugar, *explicar* un fenómeno requiere de una reconceptualización no solo del fenómeno, sino incluso del dominio en el que dicho fenómeno aparece. La redefinición conceptual del fenómeno *es parte de la esencia* de la explicación.
- El mecanismo descubierto no sólo explica el fenómeno, sino incluso *incrementa nuestro entendimiento* de dicho fenómeno.
- La explicación de los fenómenos Y1-Y4 puede derivarse de simples *tests* que pueden realizar cualquier persona por sí misma: “look and see for yourself” (2009, pág. 486)
- E1-E4 presuponen identidad entre *qualia* y procesamiento del color en el sistema nervioso. Esta identidad es meramente hipotética, no está asumida *a priori*. Las explicaciones científicas son *metafísicamente modestas*: no asumen una ontología o una metafísica *a priori*. La metafísica debe responder a la investigación empírica, no al revés. (2009, págs. 486-7)

Veamos dos objeciones relacionadas con el tipo de conexión requerida entre el explanandum y el explanans. El nagelizador propondrá que esa conexión, en un explanandum que involucra los qualia, debe ser necesaria; de otro modo se abre la posibilidad de que existan zombis: seres que satisfacen todas las condiciones del *explanans*, pero que no poseen las características propias de la experiencia, el *explanandum*:

⁵⁸ En las respuestas, incluimos incluso algunas que Foss da a otras objeciones que no mencionaremos por cuestión de espacio, pero que pueden ser incluidas como respuestas a esta misma objeción, dado que algunas de las objeciones que analiza son derivadas de esta o muy cercanas.

Objection 5: (...) To explain something is to show why it *has to be*, to reveal its necessity. But even if the neuroscientific mechanisms of colour you describe are entirely real, those who possess them may be unconscious *zombies*. Therefore, the scientific explanans does not necessitate its explanandum, so the explanation fails. (2009, pág. 489)

La respuesta de Foss ataca directamente la noción de *zombi* por carecer de contenido empírico. Los zombis son, por hipótesis, empíricamente indetectables, de modo que no tiene sentido plantearla: es científicamente irrelevante. La posibilidad de los zombis es una preocupación metafísica, lo que implica la búsqueda de algo que la ciencia no tiene por qué proveer. La conexión entre el explanans y el explanandum es de modo necesario, pero eso no implica que los principios científicos relevantes en la explicación sean por sí mismos metafísicamente necesarios: aceptamos ciertas leyes y *provisos* filosóficos, y de ellos derivamos consecuencias. Los *zombis* son *metafísicamente posibles*, pero no hay razón para preocuparse científicamente por ellos.

Pero, ¿es tan fácil librarse de esta necesidad metafísica? El explanans debe implicar la superveniencia lógica *del explanandum*, lo cual no puede satisfacerse mediante el tipo de *explanantes* E1-E4 propuestos:

Objection 7: Necessity is not so easily avoided. In scientific explanations of such things as fire or photosynthesis, the explanandum *logically supervenes* upon its microphysical substructure: in any logically possible world in which the microphysical explanans obtains, the explanandum obtains as well (Chalmers, 1996). In granting the possibility of zombies, you admit that your explanandum does not logically supervene upon its explanans, which implies that it fails by the standards of ordinary reductive explanation. (2009, pág. 491)

La apelación a la *superveniencia lógica* es una mala concepción de la explicación científica. Foss argumentará que el fuego tal como se nos aparece cotidianamente, *no superviene* a la oxidación rápida. Es posible concebir oxidación rápida sin la visión popular del fuego. Podría ocurrir que haya oxidación sin la emisión de fotones, con lo cual la visión popular del fuego no podría darse y no habría, en rigor, nada a lo que llamáramos "fuego". Así, podría concebirse a la oxidación rápida como no luminosa, con lo cual se ha eliminado el punto de vista de la superveniencia lógica en la explicación: en un caso de oxidación sin emisión de fotones, la propiedad perceptible

no supervendría a la oxidación rápida. Sin embargo, la oxidación rápida *explica* el fuego. Por lo tanto, la explicación científica no requiere de la forma de necesidad llamada superveniencia lógica. (2009, pág. 493)

Veamos, a continuación, las objeciones 9 y 10 y sus respuestas, en las que se trata de específicas propiedades de los qualia: su *esencialidad* y su *intrinsecalidad*:

Objection 9: You have explained only the accidental properties of the yellow quale, but have left its essence untouched. (2009, pág. 496)

Según esta objeción, la explicación por medio del canal de pares oponentes se detiene en el preciso instante en que debe hablar del fenómeno cualitativo.

La réplica de Foss, similar a la sugerida por Gibbons es: ¿qué puede ser más esencial a un quale que el modo en que se muestra (the way it looks)?: “Generally speaking, what makes qualia what they *are* is what they *appear* to be” (2009, pág. 496) Responde directamente a Nagel y Chalmers:

Chalmers, for instance, says that “what it *means* for a state to be phenomenal is for it to *feel* a certain way” (1996, p. 12, his emphasis), and in various places demands explanation of why seeing a colour *feels* the way it does. This literally indicates he is talking about the emotive, tactile or proprioceptive affects of seeing the colour. (...) But the affects of seeing colour is not our topic. (...) Similar comments apply to the Nagelian formulation. What is it like to see red? Maybe it is fun, maybe it is boring, maybe each time is different – but that is not the colour qualia problem. (2009, pág. 497)

Si se habla específicamente del color, puede equipararse con el ‘*way it looks*’ o ‘*what it seems*’, pero no con el ‘*feeling*’ o ‘*what is it like to be*’. Cómo se siente ver un color, o qué es ver un color no tiene que ver con el específico quale del color, sino con estados cualitativos concomitantes, simultáneos o sucesivos al quale del color. De este modo, E1 – E4 explican *lo que nos parece*₁ del color, y eso es todo lo que hay que explicar, al menos en lo que respecta a *cómo se ve* el color amarillo. Ese parecer, sin embargo, no incluye el juicio de segundo orden acerca del color (un parecer₂) o, si está presente, es irrelevante para la explicación de la experiencia de color. Que me parezca aburrido que el sol me parezca amarillo no es parte de la explicación del color.

Veamos la siguiente objeción, estrechamente relacionada con la anterior:

Objection 10: Qualia are *intrinsic* properties, but you have addressed only *extrinsic* properties, and therefore miss the mark. You proposed explanation addresses only the *relational*, hence extrinsic, properties of yellow: its brightness (its closer similarity to white than to black), its atomicity (its not being composed of other hues), its place in a colour space, and so on. You have left untouched what yellow is *in itself* (2009, pág. 497)

Foss responderá que todo concepto científico es definible de acuerdo a sus relaciones; no existen hechos que *no* posean relación con otros hechos, que es lo que se piensa cuando se postulan hechos intrínsecos. “The explanatory demand just stated would rule out all scientific explanation whatever” (2009, pág. 497) He aquí el corazón de lo que dimos en llamar “dennettización”: el nageliano demanda una explicación que violenta la noción de explicación propia de la ciencia. Su demanda estará siempre insatisfecha. Recordemos la crítica de Hofstadter a Nagel (*Supra*, pp. 168) “No quiere saber *para él* lo que es ser murciélago. Quiere saber *objetivamente* cómo es *subjetivamente* serlo (...) ¿Qué lo podría satisfacer entonces?”.

Foss prosigue atacando el propio concepto de *intrinsecalidad*:

An intrinsic property is supposed to be one that something has independently of anything else, or aside from its relationships to them. But independence itself *it's a* relationship. The absence of any given relationship is, logically, just another relationship: the fact that 3 is *not* equal 4 (the absence of equality) is just as relational as the fact that it is equal to the square root of 9 (the presence of equality). To put it paradoxically, the very concept of intrinsic property reveals them to be extrinsic. (2009, pág. 498)

Puede extraerse, incluso, la conclusión opuesta: si existe algo intrínseco, la consecuencia directa es que no puede ser conocido. La explicación científica asume las características relacionales de su objeto de estudio; de otro modo no hay manera de conectar el *explanandum* con el *explanans*.

En la conclusión del artículo, Foss resume su impresión acerca de los tipos de objeciones propuestas:

(...) [T]he most prominent arguments within contemporary philosophy of mind against the scientific explanation of consciousness, and qualia in particular, would, if correct, count against scientific explanation generally. Properly understood, these arguments are really a critique of modern science, and of part of the skeptical tradition within the

philosophy of science. Proponents of these anti-scientific views would do well to realize that it is not just psychology, or consciousness studies, that are they challenging, but science in general (2009, pág. 506)

Según la visión propuesta por Foss, el nagelizador no ataca las explicaciones científicas dadas; atacan la naturaleza misma de la ciencia. Si esto fuera así, no hay explicación científica que pudiera satisfacerlo, y esa es precisamente la impresión que había causado la cita de Nagel más arriba.

Es importante remarcar que Foss asume que la visión del color no ofrece ningún problema que requiera la apelación a leyes especiales. El hecho está suficientemente explicado hoy, y no tiene sentido pedir *algo adicional* a lo ya explicado. En todo caso, es un pedido metafísico, pero la ciencia no tiene por qué satisfacer ese tipo de demandas. Sin embargo, dejando de lado la acusación de insatisfacción psicológico – metafísica que se le achaca al nagelizador, ¿es aceptable, sin más, que los puntos E1-E4 son CAEs necesarias y suficientes, propias de la ciencia natural, para el fenómeno de la visión del color?

2.7. ¿Está Maxwell sistemáticamente equivocado en su percepción de colores?

Nuestro objetivo a lo largo de esta escrito ha sido el análisis de la legitimidad de la reducción de los estados fenomenológicos a acaecimientos explicables sin resto mediante la apelación a mecanismos puramente funcionales. El recorrido que hemos realizado a lo largo de esta parte comienza con la legitimidad del planteo fenomenológico a partir del diálogo de Dennett con Otto. Distinguimos dos tipos posibles de visión acerca del rol explicativo de la ciencia con respecto a los qualia: la visión *dennettizada* y la visión *nagelizada*. Vimos que el *parecer* cuya explicación reclama Otto es parte de una exigencia explicativa *nagelizadora*, mientras que la estrategia contraria esgrimida por Dennett consiste no sólo en *eliminar* al objeto del *parecer* de la esfera de lo que necesita ser explicado, sino incluso dejar fuera de juego a la propia noción de *parecer*. Distinguimos, en este sentido, un *parecer*₁ contenido en

el juicio fenoménico inmediato, de primer orden, de un parecer₂, contenido en el juicio de segundo orden. Aun si hubiese algo como un *parecer*₁, es posible distinguirlo de los *qualia*. John Gibbons encuentra diferencias notorias entre las características clásicas de los *qualia* y el parecer₁, y concluye con la sugerencia de que podemos quedarnos con el parecer y suspender nuestro juicio acerca de la existencia de *qualia*. Jeff Foss, tomando la sugerencia de Gibbons, explica el *what it seems* del color mediante la apelación a los canales de pares oponentes que se activan mediante rangos de estimulación de los conos de la retina. No hay fenómenos cualitativos que merezcan ser explicados una vez que se ha dado la explicación funcional: el fenómeno de la visión, una vez analizado, desaparece como tal. Parece que no queda lugar para las dudas, y todo el misterio de los *qualia* de la visión se disuelve: la manera en que nos parecen los colores puede ser explicada de modo necesario y suficiente mediante la apelación a la activación puramente funcional de canales de pares oponentes: hay identidad entre el procesamiento de color y visión (humana) de color (Foss, 2009, pág. 490) De aquí, las ilusiones ópticas, sensoriales y perceptivas pueden explicarse ya sea por la naturaleza objetiva de lo percibido, por la constitución de los sentidos o por la interpretación que realiza el cerebro a partir del estímulo: en todos los casos de ilusiones que involucran la visión de colores (el anillo *rosado* o el círculo *verde* en nuestros ejemplos anteriores), la experiencia visual (el *parecer* del color rosa en una experiencia verídica, y el *parecer* de la ilusión) dependen *directamente*, en último término, del modo de activación de los canales de pares oponentes. La reflectancia de la luz en las superficies, la longitud de onda de la luz y la actividad de los conos en la retina deben pasar por el veredicto final del proceso de pares oponentes. Si tomamos la afirmación de Foss, de que lo que ocurre en ese proceso es *idéntico* a la visión del color, podemos establecer la identidad entre el *parecer*₁ del amarillo y una determinada posición funcional en el proceso de pares oponentes. Así, las proposición “Estoy viendo amarillo” y su equivalente “*Parece*₁ amarillo” son reemplazables por “Está activo en posición positiva el canal B-Y”.

Adam Pautz (2006) enfatizará la dependencia *directa* de la fenomenología del color con respecto al proceso oponente. Las propiedades de la conciencia sensorial son tanto las que corresponden a las cualidades secundarias como a las primarias. Ambos

tipos de propiedades mentales son llamadas Q-propiedades. Pautz analizará las Q-propiedades relacionadas con el color. En la figura 10 (pág. 184), llamaremos “ φ ” a la abscisa R-G, y “ Ψ ” a la ordenada B-Y. Cuando sólo está activo uno de los canales, se tiene una experiencia de color *unitaria*. Si se activan ambos, la experiencia es *binaria*. Esta distinción entre “unitaria” y “binaria” corresponde a la distinción entre “color puro” y “color mezclado” que vimos en el apartado anterior:

Some color experiences are **unitary**. For instance, some reddish experiences are neither yellowish nor bluish. We have exactly four types of color experiences that are unitary: unitary reddish, unitary greenish, unitary yellowish, and unitary bluish color experiences. All other color experiences are binary. (2006, pág. 208)

Las nociones de “bluish”, “reddish”, etc. son parte de una terminología neutral para caracterizar la experiencia de color: “Say that a color experience is **reddish** iff it resembles the color experiences actually normally produced by objects that we call ‘red’; say that a color experience is **greenish** iff it resembles the color experiences actually normally produced by objects that we call ‘green’ and so on” (2006, pág. 208)

La propia experiencia del color posee la estructura oponente unitario – binario. Esta afirmación se basa, parcialmente, en la introspección, aunque también está sustentada por experimentos psicológicos sobre cancelación y nombramiento de colores, y se supone que recibe cierto apoyo de evidencia lingüística intercultural (2006, pág. 209)

The so called **opponent process theory** of color vision shows us that the best explanation of the character of the quality space for color experience is to be found in the brain. (...) According to opponent process theory, the values a color experience has on the phenomenal parameters red/green and yellow/blue are *determined by* the overall (relative) levels of activity in the φ and Ψ opponent channels. The opponent channels, so to speak, provide a **neural interpretation** for the quality space for color experience in the brain. (2006, pág. 209)

Las experiencias binarias resultan de la activación simultánea de ambos canales oponentes. Sin embargo, hay experiencias binarias imposibles: no existe fenomenológicamente un color que sea *rojiverde*, o *azulamarillo*, porque esto implicaría que el mismo canal se activara en positivo y negativo a la vez. La imposibilidad fenoménica de tales experiencias es paralela con la imposibilidad de activar un mismo canal *en dos posiciones diferentes de manera simultánea*.

La teoría de proceso oponente recibe cierta confirmación fisiológica aunque la evidencia todavía no es concluyente:

Direct recording of neurons has revealed that the outputs of the cones are in fact summed and differenced to create two systems of neurons in the LGN (a kind of relay station between the eyes and the visual cortex) roughly corresponding to the red-green and yellow-blue cardinal directions of color space. However, there are some discrepancies between their activity and color experience. The activity of these neurons does not match the psychophysically-derived response. For this reason and others, these systems of neurons cannot constitute the ϕ channel or the ψ channel hypothesized in order to explain the psychophysical data (...). Chromatically opponent cells are found in the different areas of the visual cortex, but even at the population level there is no evidence that they fall into distinct chromatic classes. (2006: 211)

El carácter de la experiencia de color está directamente determinado por la actividad en los canales oponentes. Esta determinación es denominada por Pautz “C-Dependencia” (2006, pág. 212). Pero esta determinación no es completa:

Note that C-Dependence says only that internal factors **play a role** in determining color experience. For this reason, it does not entail **Internalism** about color experience: the strong thesis that internal factors *completely* determines color experience, so that neurobiological duplicates living under the same laws have the same color experiences. (2006, pág. 212)

¿Por qué aceptar C-Dependencia?: “First: (...) Nothing in the outside world can explain the unitary - binary character of color experience [and] the comparative resemblance relations about color experiences. (...) So, the explication must lie in the brain. Second, systems of neurons whose activity approximates the hypothesized ϕ and ψ channels have been discovered in the early visual system.” (2006, pág. 212)

Pautz argumentará que la visión de colores depende, de manera crucial y fundamental, de factores internos. Su crítica está dirigida al *externalismo* de la percepción. El externalismo sostiene que la fenomenología de las experiencias visuales representa propiedades externas. Dos sujetos pueden conocer de modo idéntico la propiedad externa Φ aun cuando sus representaciones internas difieran. A su vez, un cambio en las propiedades externas provocará experiencias diferentes. Pautz argumentará que dos individuos que posean un cableado posreceptorial diferente, tendrán diferentes experiencias aun cuando las propiedades externas se mantengan sin diferencia (2006, pág. 213)

Supongamos un mundo posible muy cercano al nuestro, *W*, cuyas leyes fundamentales son idénticas a las de nuestro mundo, y en el cual se siguen las dos condiciones siguientes:

- a) Sus individuos humanos poseen una historia evolutiva diferente a la nuestra: las curvas de respuesta de sus canales de pares oponentes *difiere* con respecto a nuestro mundo.
- b) Las curvas de respuesta de los tres tipos de conos de la retina se mantienen sin cambio.

En consecuencia, dadas las mismas condiciones de reflectancia de luz, de activación sensorial y de respuesta de los conos de la retina, la experiencia de los habitantes de *W* será diferente a la de los habitantes de nuestro mundo, en virtud de la diferencia en el cableado de los canales de pares oponentes.

Supongamos que en *W* los mapeos de la reflectancia de luz no conservan la distinción unitario - binario propia de nuestro mundo. Por caso, si un individuo de nuestro mundo tiene una experiencia unitaria con la propiedad de reflectancia *R*, un individuo en *W* tendrá una experiencia binaria y viceversa. El veredicto correcto es: *mismas causas externas, pero diferente procesamiento dan como resultado diferentes experiencias*. Esta diferencia en las experiencias dará como resultado una disposición a conductas *diferentes*. Si tomamos un par de individuos, Maxwell (el habitante de nuestro mundo) y el Maxwell Gemelo (habitante de *W* idéntico a Maxwell excepto por la reconfiguración del cableado de pares oponentes), quienes son estimulados por el mismo espectro de reflectancia luminosa *F&H*, podemos diagramar la situación de la siguiente manera (Figura 11) (2006, pág. 215-6):

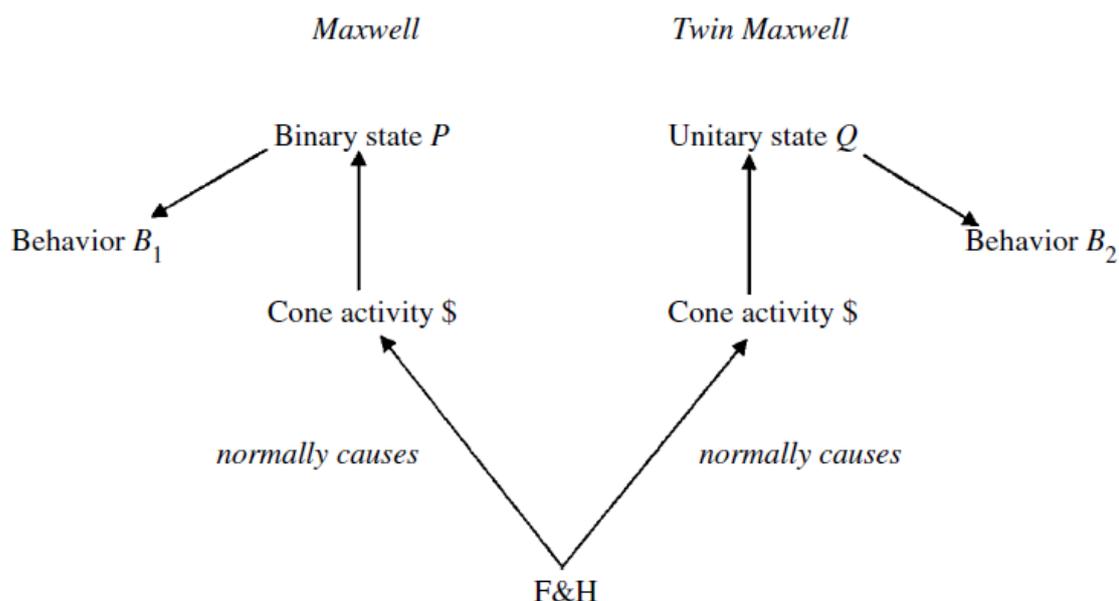


Figura 11. Diferencias entre Maxwell y Maxwell gemelo. (Extraído de Pautz, 2006, págs. 215 – 216)

Los externalistas son fisicalistas en lo que respecta al color: esto es, el color se identifica con sus propiedades de reflectancia. Por lo tanto, el fisicalista explica la distinción unitario – binario en términos físicos. Una indeseada consecuencia de este punto de vista es que nuestra imposibilidad de ver o imaginar colores como *rojiverde* o *azulamarillo* se debe a que, en términos físicos, esos colores simplemente *no se dan en la naturaleza*, lo cual resulta extremadamente implausible. Otra consecuencia es que los colores son binarios o unitarios por causas de la propia reflectancia de la luz, con lo cual se afirmarían que los canales de pares oponentes juegan un rol irrelevante, o ninguno en absoluto:

On externalism, th[e] reflectance property is necessarily identical with a binary [or unitary] color. But if opponent channel states play a **direct** role in determining the unitary-binary character of color experience, in accordance with C-Dependence, then this should be irrelevant. (2006, pág. 218)

Maxwell y Maxwell gemelo, a despecho de la inversión unitario – binario que poseen, deberán tener, según el externalismo, *mismas experiencias*, dado que el estímulo es el mismo. Si la propiedad de reflectancia es determinante para tener una experiencia

unitaria o binaria, entonces la inversión en la unitariedad o binariedad de los canales de pares oponentes no jugará ningún rol.

Debe tenerse en cuenta, además, que tanto Maxwell como Maxwell gemelo están en *condiciones perceptivas óptimas*: la diferencia en los procesos evolutivos de ambos no ha generado una desventaja adaptativa en uno o en otro. Que ambos estén en condiciones óptimas de percepción es la *tesis de la simetría* entre el Maxwell y Maxwell gemelo: ambos individuos poseen *igual capacidad adaptativa*. (2006, pág. 223)

Podría cuestionarse que, en el ejemplo imaginado, se ha forzado una situación en la cual se decide *a priori* que Maxwell gemelo tiene condiciones perceptivas óptimas: podría argumentarse que *si* un individuo posee canales que le otorgan experiencias binarias cuando en nuestro mundo dichas experiencias son unitarias, *entonces* su experiencia es producto de un mal funcionamiento. ¿Es posible imaginar un mundo cercano en el que, por ejemplo, los seres humanos evolucionaron de modo que sólo tengan la experiencia visual del color blanco, y a su vez que dicha evolución es óptimamente adaptativa? Se puede imaginar lo primero, pero difícilmente lo segundo. Sabemos que hay disposiciones óptimamente adaptativas a partir de los datos de *nuestro* mundo; no podemos decidirlo *a priori* para otros mundos posibles.

Sin embargo, este argumento no provoca dificultades. No podemos pensar que en *nuestro mundo* hemos sido dotados del *único* conjunto óptimo de disposiciones. Si así fuera, debemos pensar que la evolución nunca ha estado en estado de condiciones adaptativas óptimas hasta el día de hoy: hemos tenido una suerte extraordinaria (*The Grand Fluke Thesis*). En vista de estas objeciones, resulta más plausible aceptar que existen multitud de procesos evolutivos óptimos en las mismas condiciones ambientales. De allí, es posible aceptar que Maxwell y Maxwell gemelo, a despecho de su diferencia evolutiva y su idéntico entorno actual, puedan estar máximamente adaptados a su entorno. (2006, pág. 225)

La conclusión de Pautz, reforzada en Pautz (2010), es que la experiencia de color no puede reducirse a un proceso físico funcional relacionado con las propiedades de reflectancia luminosa: la experiencia del color es, en un sentido importante, **primitiva**,

es decir, parcialmente independiente de lo que ocurre en el mundo exterior y parcialmente independiente de los procesos funcionales subyacentes.

La conclusión de Pautz difiere de la de Jeff Foss. Mientras para este último hay *identidad* entre el procesamiento de color y visión de color, para Pautz hay *dependencia* entre la experiencia y el procesamiento, aunque no *identidad*. Las propiedades de la experiencia no pueden identificarse con procesos físicos – de allí el *primitivismo*, y de aquí la irreducibilidad de la experiencia a los procesos físico – funcionales subyacentes. En un sentido importante, sin embargo, ambas conclusiones se asemejan. La explicación de la experiencia de color mediante pares oponentes no deja de ser una explicación funcionalista que no parece superar el gap explicativo entre la experiencia y sus procesos subyacentes, en tanto el sistema de pares oponentes es un esquema que tiene semejanza con un proceso neuronal. Pautz enfatiza que *los factores internos juegan un rol en la determinación de la experiencia*, (207) y a su vez que la explicación de la distinción unitario – binario debe residir *en el cerebro* (212). Si por “factores internos” Pautz entiende procesos neurales del cerebro, entonces es difícil entender en qué sentido las Q- propiedades son primitivas, dado que “no pueden identificarse con ninguna región corporal” (207), pero sin embargo dependen **directamente** de procesos neurales. Parece que cuando Pautz dice “factores internos” se refiere, efectivamente, a algo que no depende ni de la reflectancia de la luz, ni del funcionamiento de los sentidos ni de los conos de la retina, sino a procesos neurales. Pero estos procesos son tan *externos* como el resto, al menos desde la propia fenomenología de las Q-propiedades. Una consecuencia indeseada de este primitivismo es que no explica la relación con los procesos neurales y sólo se limita a señalar su dependencia *directa*, sin dar mayores detalles acerca de ese calificativo. Tanto para Pautz como para Foss, la teoría de los canales oponentes es el principal foco de atención para analizar las correspondencias entre las Q-propiedades y los procesos neurales.

Alex Byrne y Michael Tye (2006) atacarán los principales argumentos de Pautz. La tesis de la C-Dependencia será puesta en duda, en primer lugar, sobre la ausencia de datos concluyentes acerca de la existencia de algún proceso funcional en el cerebro que se comporte tal como se esquematiza en la teoría de los canales oponentes (2006, pág.

243). En segundo lugar será cuestionada la relación entre el proceso de canales oponentes y los estados cualitativos del color. Pautz dice que la actividad en los canales oponentes *determina directamente* el carácter de la experiencia resultante.

Byrne & Tye replican:

What does ‘determines’ mean in the statement of C-Dependence? Pautz tell us that “[t]he modally weakest interpretation of C-Dependence is a Dualist one... [on which] C-Dependence is supported by a brute psychophysical law directly linking opponent activity with color experience” (212). On the other, “Physicalist”, interpretation of C-Dependence, “color experiences are somehow constituted by opponent channel states” (217). All this strongly suggests that ‘determines’ means *either nomologically or metaphysically necessitates*.

What about ‘directly’, which modifies ‘determines’ in the statement of C-Dependence? (...) Although Pautz bold-faces the word, and so presumably thinks it cannot be removed without loss, it is unclear what he has in mind. One possibility is this: proposition P *directly* determines proposition Q iff P determines Q and P does not determine any other proposition that determines Q. But this can hardly be right, because if P and Q are contingent and distinct (as they will be in cases of interest), then P will *not* directly determine Q. In any case, as far as we can see, ‘directly’ is not doing any work in Pautz’s argument, so we will ignore it in what follows. (Byrne & Tye, 2006, pág. 244, cursivas en el original)

Byrne y Tye desestiman el uso de “directamente”, cuando Pautz afirma –incluso subrayando la palabra– que la C-Dependencia determina *directamente* la experiencia del color. La determinación –sin calificativo– puede entenderse como la necesidad nomológica o metafísica de que, dada la posición positiva en el canal φ , se obtendrá inmediata y necesariamente la experiencia de rojo. Byrne & Tye, entonces, reformulan la tesis de la C-Dependencia, reinterpretándola en términos de una C-Dependencia fuerte:

Strong C-Dependence: internal neural states and processes (specifically opponent channel internal states and processes) either nomologically or metaphysically necessitate the character (qualia) of color experience. (Byrne & Tye, 2006, pág. 244)

Sin embargo, la tesis de Pautz no presupone que los procesos internos del canal oponente determinen *completamente* la experiencia del color. Debe recalarse que la aceptación de la tesis fuerte de la Dependencia implicaría que en dos individuos que sean duplicados físicos exactos, será nomológica o metafísicamente imposible que tengan diferencias en sus experiencias de color, *sin importar el estímulo externo*. La

exacta duplicación de los factores internos implica una idéntica experiencia. Sin embargo, cuando Pautz, al principio de su artículo, explica en qué consiste la *Dependencia*, la tesis general de la cual la C-Dependencia es una especie, escribe acerca de casos “en los que dos posibles individuos estén en diferentes estados neurales internos”. En esos casos:

Dependence is *no stronger* than the claim that, *at least in some such cases*, the correct verdict is *Different Experiences*: the individuals involved have different experiences. (207, our emphasis) (Byrne & Tye, 2006, pág. 245)

De acuerdo a esta aclaración, la C-Dependencia, como caso especial de la dependencia, no hace necesario, de manera nomológica o metafísica, la aparición de una determinada experiencia cualitativa. Por lo tanto, la tesis de C-Dependencia no es fuerte; puede ser entendida en una versión debilitada:

Weak C-Dependence: color experiences are not just *correlated* with internal neural states and processes (specifically opponent channel internal states and processes). Suppose that opponent Channel state O is correlated with quale Q: a perceiver’s opponent channels are in O iff she is having an experience with quale Q. Then: if a perceiver’s opponent channels *had been* in state O, the perceiver *would have* had an experience with quale Q. (Byrne & Tye, 2006, pág. 245)

En esta versión debilitada, el antecedente del contrafáctico no es una ley, sino un enunciado acerca de hechos individuales. Pautz da dos argumentos para aceptar la C-Dependencia. El primero es:

Nothing in the outside world can explain the unitary-binary character of color experience, the comparative resemblance relations among color experiences (...). So, the explanation must lie in the brain. (Pautz, 2006, pág. 212) (Byrne & Tye, 2006, pág. 246)

Si se interpreta esta afirmación como un argumento para la C-Dependencia fuerte, se estaría afirmando que el carácter unitario – binario de la experiencia puede ser *completamente* explicado por hechos internos. Sin embargo, Pautz no excluye la posibilidad de que los factores externos sean *parte* de la explicación del carácter unitario-binario de la experiencia del color (2006, pág. 246). El problema es cómo compatibilizar la afirmación “*nothing in the outside world can explain...*” con la idea de que *parte* de la explicación puede invocar factores externos, que es lo que permite la afirmación de la C-Dependencia debilitada.

Pautz da un segundo argumento para aceptar la C-Dependencia:

Why accept C-Dependence?... Systems of neurons whose activity approximates the hypothesized [opponent] channels have been discovered in the early visual system (112) (Byrne & Tye, 2006, pág. 247)

Byrne y Tye remarcan que este argumento no funciona, dado que el propio Pautz ha afirmado previamente que “these (LGN) neurons ‘cannot constitute’ the hypothesized opponent channels (212)” (Byrne & Tye, 2006, pág. 247), pero aun si de hecho las neuronas constituyeran los canales oponentes, esto no es suficiente para afirmar la C-Dependencia fuerte. Un descubrimiento de las neuronas que se comportan tal como lo describe el sistema de pares oponentes no proporcionaría el *locus* de la *experiencia* de los colores, sino la *tipografía neuronal* por medio de la cual los colores pueden ser experimentados, del mismo modo que los tipos de un directorio telefónico no pueden identificarse con los números de teléfono de las personas. “Typography does not determine semantics” (Byrne & Tye, 2006, pág. 247). Nos queda, entonces, la C-Dependencia débil. ¿Permite inferir esta tesis que Maxwell y Maxwell gemelo poseerán diferentes experiencias? Dicho de otro modo: en un mundo posible en el cual las leyes evolutivas difieren levemente en el cableado de los canales oponentes mientras todo lo demás se mantiene constante, ¿se puede concluir que ambos tendrán experiencias diferentes *en virtud de su diferencia de cableado, aun cuando los estímulos externos se mantuvieran idénticos en ambos mundos?*

The argument from weak C-Dependence for Different Experiences may be set as follows:

P1 (From weak C-Dependence). Maxwell is looking at an orange square. His opponent channels are in state O, and he is having an experience with quale Q. If his opponent channels had been in a different state O*, he would have had an experience with a different quale Q*.

P2. Twin Maxwell (who, for simplicity, can be identified with Maxwell himself) is in a “nearby counterfactual situation” (213) in which evolution goes slightly differently, and in which the causal connections between external conditions and Twin Maxwell’s internal states are slightly different. Twin Maxwell is looking at a square of the very same color, and his opponent channels are in state O*.

Hence, (appealing to implicit background details about the actual and counterfactual situations):

C. Twin Maxwell’s experience (in the counterfactual situation) has Q*, and so Maxwell (in the actual situation) and Twin Maxwell are having different experiences. (Byrne & Tye, 2006, pág. 248)

Según Byrne y Tye, el argumento es inválido, porque en la base implícita se está presuponiendo que el mundo de Maxwell Gemelo es *el más cercano* a nuestro mundo. Pero en el mundo más cercano no puede aceptarse, como Pautz propone, que la evolución humana ha ido por carriles levemente diferentes. “The closest world must keep evolution fixed” (Byrne & Tye, 2006, pág. 248)

En el ejemplo de Pautz, tanto Maxwell como Maxwell gemelo están equipados con un conjunto óptimo de disposiciones adaptativas y por lo tanto, no puede decirse que uno de ellos posea percepciones erróneas. Byrne y Tye se preguntan, con respecto a esta afirmación, cómo es posible que *en óptimas condiciones*, dos individuos puedan haber evolucionado de maneras tan diferentes –si se sostiene, desde luego, que se trata de un *mundo cercano* al actual. En un mundo así, todo el sistema de relaciones de colores de Maxwell gemelo será diferente. Tal sistema de percepción covariaría con el camino evolutivo que se ha seguido en ese mundo. Teniendo en cuenta esto, es difícil predecir si su experiencia será *diferente* o *similar* a la de Maxwell. Quizás, en ese mundo, el recableado provoque que, cuando en nuestro mundo se activa uno solo de los canales oponentes, en ese mundo se activen ambos. Pero quizás en ese mundo, en ciertos casos *la activación de dos canales provoque una experiencia unitaria y no binaria* como pretende Pautz. Por lo tanto, el veredicto de Pautz, de que Maxwell y Maxwell gemelo tendrán diferentes experiencias sobre la base de la diferencia en los pares oponentes no puede seguirse inmediatamente de las estipulaciones dadas en su ejemplo. (Byrne & Tye, 2006, pág. 253)

Es importante remarcar que el artículo de Byrne y Tye no ataca el *primitivismo* de las experiencias conscientes, sino la legitimidad de la apelación a la teoría de pares oponentes para explicar las experiencias. Al principio del artículo, Byrne y Tye citan a J. D. Mollon (1997) quien resumía la situación de la identificación o la relación entre los canales oponentes y la experiencia del color:

We still believe today that there are chromatically antagonistic channels in the early visual system—that is, channels that draw inputs of opposite sign from different classes of cone, but these channels simply do not correspond to the phenomenologically defined channels of Hering. . . . In fact, no one has found a site in the visual system where colour appears to be represented according to Opponent Colour theory—that is, a site where the cells might be held to secrete

redness and greenness or yellowness and blueness. Cells are found in the cortex that respond to restricted regions of chromaticity space, but they are by no means confined to the *loci* of pure hues. Thirty years ago we thought we understood the existence of four unique hues, hues that are phenomenally unmixed. Today this is perhaps the major unsolved problem of colour vision (Mollon & Jordan, 1997). If we understood it, we should probably be much closer to understanding the general relationship between neural activity and qualia. (Mollon 1997, págs. 870–872. En Byrne & Tye, 2006, pág. 243)

Según Mollon, el proceso oponente correspondiente a los colores no tiene una corroborada correspondencia con la actividad de neuronas en la corteza cerebral. Si esto es así, entonces toda hipótesis que utilice la teoría de los pares oponentes para explicar tanto la sensación de color como la *imposibilidad* de algunas de ellas (rojiverde y azulamarillo), no habrá hecho ningún avance definitivo en la explicación del *gap* entre los estados cualitativos de conciencia y el sistema funcional que le sirve de soporte. En principio, la afirmación de Pautz de que la explicación del carácter del espacio cualitativo para la experiencia del color debe encontrarse en el cerebro es, por lo menos, infundamentada. Él mismo reconoce que los resultados de la asociación entre experiencia, pares oponentes y semejanzas funcionales con procesos neuronales son problemáticos e inconcluyentes⁵⁹.

Pero quizás el punto más importante es que, tal como se desprende de la cita de Mollon, la propia teoría de los canales oponentes es, ella misma, puramente *fenomenológica* y data del esquema de las experiencias de color de E. Hering, (1878):

In 1878 Hering had proposed that our colour experience depended on two neural processes, each bipolar: an antagonistic red-green process and an antagonistic yellow – blue process. The evidence then was phenomenological and, I shall to argue, it is still the only evidence today. (1997, págs. 867-8, cursivas añadidas)

Dicha “evidencia fenomenológica” pareció poder articularse con el funcionamiento de grupos de neuronas, a partir de los estudios de De Valois durante los años ’60 efectuados en macacos, aunque tal articulación es, según Mollon, débil y apresurada:

⁵⁹ Ver infra: Pautz, 211: “there are some discrepancies between their activity and color experience. The activity of these neurons does not match the psychophysically-derived response. For this reason and others, these systems of neurons cannot constitute the ϕ channel or the ψ channel hypothesized in order to explain the psychophysical data”

What was taken as evidence for Hering's theory was the physiological work of De Valois and his collaborators (De Valois, 1965; De Valois, Abramov, & Jacobs, 1967). Recording from single units in the lateral geniculate nucleus of macaques, they found cells that were excited by one part of the spectrum and inhibited by others wavelengths. Such units appeared to fall into four classes and were quickly labeled with the initials of colour names: +R-G, +G-R, +B-Y and +Y-B. These findings led very sidely to the idea that Hering had been vindicated and to the statement that the Young-Helmholtz trichromatic theory held at the level of the receptors while the Hering Opponent Colour theory held at a postreceptoral level. You still find this line in textbooks, and I know that it is taught to this day in more than one British university. Here is a case where psychologists were misled by physiology. (1997, pág. 868)

Recordemos que tanto Foss como Pautz recurren a la teoría de los canales oponentes y que tal teoría pretende explicar el funcionamiento del sistema posreceptoral que es responsable funcional *directo* de nuestras experiencias de color. Una de las intenciones de Jeff Foss era reducir la explicación acerca de la fenomenología del color a términos acerca de pares oponentes. Sin embargo, si esta última teoría es puramente fenomenológica, entonces se ha explicado la fenomenología con más fenomenología. Foss argüía que la explicación de los qualia de color exigía una reconceptualización del fenómeno a explicar, y que, en el caso del color, tal reconceptualización era un trabajo de rutina (*humdrum*) desde un punto de vista científico: no se requería de un principio innovador para explicarlo (Foss, 2009, pág. 484). Sin embargo, si se ha tomado la teoría de canales oponentes como instrumento reconceptualizador, la explicación falla, puesto que tal reconceptualización está dentro del terreno de las propias experiencias fenoménicas que se deseaban reconceptualizar. La explicación E1-E4 apela de manera directa a la teoría de los canales oponentes y, de ese modo, de acuerdo a la objeción de Mollon, no incorpora una auténtica CAE de ciencia natural. Retomemos parte de la explicación E4 dada por Foss más arriba en § 2.6: "Since yellow results from an opponent process with blue, it cannot blend with blue" (Foss, 2009, pág. 481). La explicación de la química del color amarillo; esto es, la capacidad de formar compuestos con algunos colores, pero no con otros; por ejemplo, la imposibilidad de un color *azulamarillento*, está dada por el *locus* oponente del color amarillo y el azul. Pero la oposición entre amarillo y azul en la teoría de pares oponentes es ella misma una derivación del *parecer*₁, esto es, producto de la fenomenología humana ("No puedo siquiera imaginar el azulamarillo") y por lo tanto no cuenta como explicación

fisiológica de esa fenomenología; es una mera tautología decir que el amarillo no puede mezclarse con el azul porque de acuerdo a nuestro parecer tal mezcla es imposible. La denetización de la explicación no se ha logrado en absoluto y quizás Foss, sin proponérselo, haya caído en la trampa de la “fenomenología objetiva” que reclamaba Nagel.

2.8. ¿Sueñan las palomas con direcciones aéreas ultravioletas?

En la primera parte de este trabajo establecimos que el mecanicismo clásico no podía realizar una reducción satisfactoria entre la fenomenología (las ideas derivadas de cualidades primarias y secundarias), y el mundo externo puramente mecánico (de donde se derivan las cualidades “reales”). A partir de las confrontaciones expuestas en los dos últimos párrafos, encontramos que algunos de los intentos contemporáneos de reducción de fenómenos cualitativos de la visión del color a sus componentes funcionales tampoco podía lograrse, debido a que en los explanantes estaban involucrados elementos fenomenológicos. En ambos casos, separados por dos siglos y por una gruesa cortina conceptual (la idea de los modernos difiere sustancialmente de los *qualia* o el parecer₁), la fenomenología parece refractaria a los sucesivos tratamientos reductivos. En ambos casos, tanto en la modernidad como en los intentos contemporáneos examinados, se toma como base la fenomenología *humana*. Tal vez, uno de los problemas con respecto a la reducción estriba en que necesitamos una idea fenoménica general, que no se aplique sólo a seres humanos sino a cualquier ser que perciba. Es tal vez la falta de atención a esta idea general la que nos lleve por mal camino: cuando se establece que los pares oponentes son necesarios para explicar en qué consiste la visión del color, se está haciendo referencia a un marco fenomenológico puramente humano. Pero otras especies tienen sistemas de relaciones fenoménicas que difieren de los humanos. En este apartado examinaremos los alcances de una definición general de “experiencia visual”.

Cuando Foss propone su explicación del *qualia* de color amarillo, lo hace en los términos de Gibbons: *why colours look the way they do* (2009, pág. 480) Foss no

pretende explicar qué es el amarillo definido como color cuyo rango se encuentra entre 565 y 580nm, sino cómo vemos o cómo nos *parece*₁ la *visión* de la luz en (algún valor de) dicho rango. Las definiciones *folk* se basan parcialmente en diferencias fenoménicas entre los objetos que paradigmáticamente nos provocan el *parecer*₁ de ciertos colores: “El amarillo es el color del sol”; “El naranja es el color de las mandarinas maduras”, etcétera. Tales objetos nos sirven para fijar una referencia que pueda ostentarse de modo *externo*. No es una sorpresa que los nombres de los colores a veces se identifiquen con su objeto paradigmático (el naranja con las naranjas, el celeste con el cielo). Al menos en su sentido *folk*, los colores son colores *de* algo que *no está* en la mente, excepto en casos marginales. Cuando decimos “El oro es amarillo” quizás esta expresión sea identificable con “El oro, en presencia de una fuente luminosa como el sol, refracta una luz cuya longitud de onda está entre los 565 y 580nm”. Pero cuando decimos “Me parece₁ que el oro es amarillo”, la identificación con un proceso óptico cuantificable es más discutible: la *experiencia visual* del amarillo del oro no nos *parece*₁ una longitud de onda determinada: simplemente nos parece amarillo en virtud de un determinado tipo de representación espontánea e inmediata que es independiente de la efectiva longitud de onda de la luz que refleja en el objeto. La explicación científica de los qualia debería poder soportar contrafácticos del tipo “Si y solo si X hubiera estado en presencia de una fuente luminosa cuya longitud de onda está entre los 565 y 580nm, entonces le habría parecido que tenía una experiencia visual del color amarillo en el objeto Y”. El problema es que esta identificación sólo es válida si X es un individuo humano normal (dejando de lado, por ahora, las restricciones necesarias para hablar de “normal” en este caso). Pero, ¿puede el color, en tanto es experimentado, definirse de acuerdo a nuestro *parecer*₁ humano? ¿Podemos sostener este contrafáctico si X es un individuo de otra especie que también puede ver color, como es el caso de las palomas? Parece que la definición de color en términos de *experiencia visual* debería incluir tales casos. Mohan Matthen, en su artículo *The disunity of colour* (1999), desenmascara el *chauvinismo* presente en las definiciones de color en términos de experiencia humana:

Philosophers have standardly defined color in terms of human experience. (...) Let’s look at a typical anthropocentric definition. David Lewis has recently proposed the following as a schema for defining color terms:

D1. *Red* is the surface property which typically causes experience of red in people who have such things before their eyes.

D2: *Experience of red* is the inner state of people which is the typical effect of having red things before the eyes.⁶⁰

(...) Lewis avers that colors are *reflectance* properties of surfaces. So to get a general conception of color out of his definitions, we need to amend D1 accordingly: "*Red* is the *surface spectral reflectance* which typically causes...etc."

Consider, then, the following formulae:

S1 *x* is the surface spectral reflectance which typically causes *experience of x* in people who have such things before their eyes.

S2 *Experience of x* is the inner state of people which is the typical effect of having *x* things before the eyes. (Matthen, 1999, pág. 48)

Podemos sustituir *Red* por cualquier otro término de color. La experiencia de color requiere que se dé simultáneamente el par S1 + S2, pero las condiciones puestas en ambas tesis parecen particularmente restrictivas: S1 sólo puede darse ante *personas*, y S2 sólo puede ser la experiencia de *personas*. Si existe una propiedad de reflectancia de una superficie (tal como exige S1) pero esa propiedad de reflectancia no provoca una experiencia en individuos *humanos*, entonces esa propiedad no es un color:

Imagine an organism, *O*, [...] with its short wave receptor shifted toward the ultraviolet. It will be capable of seeing ultraviolet reflectances. (...) [I]t would experience ultraviolet in much the same way as we see violets and purples (that is, a bluish-red). Ultraviolet is not a color according to S1: spectral reflectances in this range do not "typically" cause "people" to have any sort of experience at all, even when they are "before the eyes". However, S2 admits *O*'s experience of ultraviolet; it is the very same "inner state" which, in people, is "the typical effect of having violet things before the eyes." Lewis is thus obliged to say that although *O* experiences ultraviolet as a color, it is not a color. This odd result, which seems incongruent with the general idea of pairing colors with kinds of experience, shows how his definition turns on contingent features of human color experience. (1999, pág. 50)

Las palomas poseen fotorreceptores similares a los de los humanos, pero poseen, además, un pigmento sensible al ultravioleta. Su experiencia visual no es describable en términos de los componentes de la experiencia humana. La definición de Lewis recorta la experiencia de color únicamente a la *humana* y, en ese sentido, parece arbitraria:

⁶⁰ Lewis, (1997) Citado en Matthen (1999), pág. 48.

Conclusion 1: Anthropocentric accounts that restrict colors to those detected by humans, or color experiences to those that occur in humans, cannot account for color vision as it occurs in other species. (1999, pág. 51)

Pero el antropocentrismo podría aceptarse *al menos* como punto de partida de una definición de color. Quizás, como dice C. L. Hardin (Hardin, 1988),

We can begin with the human case, in which we know there to be genuine color vision, and extend the concept of a color perceiver outward to other species. (Hardin, 1988. En: Matthen, 1999, pág. 51)

¿Por qué sería necesario partir del caso humano? Porque la *experiencia* de color no puede definirse *funcionalmente*, es decir, como una capacidad de respuesta diferencial a las longitudes de onda. Si estamos hablando de la experiencia del color, nos estamos refiriendo a la fenomenología, al parecer₁. Es necesario que en la visión de color haya algún tipo de representación interna y, por ello, nuestro propio caso provee un ejemplo paradigmático de este tipo. Puede construirse un sistema consistente en un receptor sensitivo de longitudes luminosas de onda larga, conectado a un motor que lo hace mover de acuerdo a la fuente de luz. Pero difícilmente diríamos que ese sistema ve los colores. Por ello, el punto de partida de la definición de *visión de color* debería incluir a individuos que son capaces de cierta representación interna del color. Las categorías fenoménicas básicas de las que se debe partir son las tonalidades que percibimos los humanos: el habitual parecer₁ del rojo, verde, amarillo y azul. En definitiva, nuestro propio procesamiento oponente nos provee de un punto de partida para elaborar una definición más amplia.

Sin embargo, nuestro sistema visual depende de dos subsistemas: un sistema filogenéticamente reciente revestido sobre otro más antiguo:

The ancient subsystem (...) is dichromatic, and it detects a very simple characteristic of the wavelength distribution of a spectral signal. (1999, pág. 55)

Este subsistema divide nuestras sensaciones de color en *cálidas, frías o neutrales*.

Matthen propone que imaginemos una especie que posee el subsistema antiguo, pero no la división de tonos propia de los canales oponentes. ¿Diríamos que ve color?

(...) [D]oes it have color vision? On grounds of phenomenology, and in isolation, one might well be inclined to say no: in our color phenomenology, warm and cool seem to

be extrachromatic characterizations that sit on top, as it were, of the colors. But when we remember that warmth and coolness is pretty much the same as the long/short wavelength determination that forms the basis of our blue-yellow categories, we should hesitate. (1999, pág. 55)

Tanto la distinción “cálido – frío”, como “azul- amarillo” dependen de propiedades de la longitud de onda luminosa:

(...) [W]hen we remember that warmth and coolness is pretty much the same as the long/short categories, we should hesitate. Clearly the primordial subsystem has access to one of the two dimensions of our richer experience of hue. To deny that it represents the world in color seems as presumptuous as for a tetrachromat to sniff at us. (1999, pág. 55)

La experiencia de visión dicromática (cálido – frío) es *una parte* de la experiencia de color. Un animal que posea el subsistema antiguo, pero que no posea el más moderno, tendrá, con todo, experiencias de color. La particular fenomenología humana con respecto a la visión de color no puede entenderse como la base para el estudio de la visión *en general*. “Why should we assume that the human system is the norm rather than some kind of oddity or specialized adaptation?” (1999, pág. 58). Puede derivarse una segunda conclusión:

Conclusion 2: Defining *color vision* by its overall *resemblance* to the human case falls short of articulating *explanatory* principles that govern the classification of color-like experiences in other species (1999, pág. 58)

Recordemos que, por un lado, necesitamos definir la visión del color de acuerdo a un sujeto percipiente que pueda tener algún tipo de representación interna. De otro modo, sólo nos quedamos con una definición puramente funcional del color que no implica necesariamente una experiencia (un parecer₁) de color. Pero resulta demasiado restrictiva una definición para la cual la visión humana es un caso paradigmático. Por ello, quizás el problema está en definir exactamente en qué consiste el *color* más que en la *visión* de dicho color. Hilbert (1992, pág. 364. Citado en Matthen, pág. 59) propone el “realismo antropocentrista”:

(a) [I]t is possible to identify our own color experiences introspectively as “an aspect of our visual experience that is clearly different from figure, texture, depth”. He suggest (b) that the human color experiences so identified have the biological function of detecting surface spectral reflectances. This, he thinks, makes it plausible (c) that colors are surface spectral reflectances. (1999, pág. 59)

La capacidad de detectar introspectivamente reflectancias espectrales en las superficies permite que otras especies que detectan reflectancias que no están en nuestro espectro de visión, también sean incluidas dentro de la definición de “visión de color”. La detección del color ultravioleta en las plumas de las aves es un caso de visión de color, dado que provoca una experiencia de color que detecta reflectancia. El ultravioleta es, de acuerdo a la definición de Hilbert, un color.

Pero (c), la definición funcional de color, provoca algunas consecuencias contraintuitivas. ¿Es suficiente la definición de color como la reflectancia luminosa en las *superficies*?

Let's look again at the pigeon's ultraviolet vision. The atmosphere scatters ultraviolet light more than light of longer wavelengths. (...) [T]he pigeon is able to navigate in an aerial environment which not only lacks “landmarks” but extends deep into three dimensions (by contrast with our own environment which does not normally extend far below our feet or above our eyes). Even in trackless space, the bird can determine direction in heliocentric coordinates. (...) The pigeon detects a property of direction, and direction is not a surface.

(...) We discriminate surfaces and locate them by the color experiences induced in us by light differentially reflected from them. This is at least part of our justification for saying that surfaces look colored. The bird discriminates *directions* by the color-like experiences induced in it by light differentially scattered from them. This should incline us to say that to the bird, directions look colored: the direction of the sun looks white, and directions at right angles to the sun looks ultraviolet. (1999, pág. 60)

Pueden verse las limitaciones del realismo antropocentrista: que la reflectancia de la luz en las superficies sea lo que normalmente provoca en humanos la visión del color, no implica que las superficies sean la *única* causa externa de tales experiencias: las direcciones de la luz en un espacio (y no en una superficie) también provocan experiencias de color. Incluso en nuestro propio caso hay al menos una experiencia de color que no está provocada por una superficie: el azul del cielo. “We say that the sky is blue, and it is not a surface” (1999, pág. 61). Matthen aclara que, de todas maneras, en el caso humano el azul del cielo parece ser una “mala percepción normal”, un subproducto del modo en que el sistema visual humano procesa el color en las superficies. El de las palomas no es el mismo caso: ellas perciben el color de las *direcciones* de la luz solar.

Parece, entonces, que la *reflectancia en las superficies* no es suficiente para definir funcionalmente a *color*. El objetivismo requiere de alguna propiedad que sea *independiente* de lo mental, para poder determinar si un organismo posee visión de color o no la posee. En ausencia de tal propiedad, únicamente contamos con el caso humano como base. Pero, como señala Matthen:

[I]t is far from clear why the objectivist must insist on there being a *single* mind-independent property that is the color. Why not allow for a plurality of mind-independent properties? (1999, pág. 61)

El problema con la aceptación de una pluralidad de propiedades independientes de lo mental es que se necesita explicar la *unidad* de la experiencia de color. Las reflectancias espectrales son *reales* y pueden ser dispuestas en un espacio cualitativo medido numéricamente: “Hilbert looks to this quality space to provide him with a real counterpart of the unity of experienced color” (1999, pág. 62) Las propiedades de las direcciones no forman parte del espacio de reflectancia, y por lo tanto deben ser desplegadas en un espacio cualitativo diferente e inconmensurable con el espacio de la reflectancia en superficies. Pero,

[t]he trouble with this proposal is that it leans too heavily on the claim that reflectance is the one and only thing that human color vision functions to detect. What we have learned from the case of the pigeon is that the mechanisms of color vision *can* be used to detect other things. (1999, pág. 63)

Matthen incluye, dentro de la propia visión de color humana, la detección de otras propiedades además de la reflectancia: *luminancias* y *transmitancias* (luminances and transmittances)⁶¹. “Indeed, we are more like the pigeons than Hilbert is ready to acknowledge. Not all of our colors are reflectances, not all are properties of surface. Not every color of which we are aware is even *located* – the blue of the sky has a direction (like pigeon’s ultraviolet?) but no obvious location.” (1999, pág. 64) Si aun en el caso humano, la experiencia de color detecta un sinnúmero de propiedades, no es posible asegurar que una *dirección* no puede poseer color. De aquí:

Conclusion 3: Anthropocentric definitions of *color* in terms of the alleged real unity of colors as detected by humans cannot accommodate the diversity of properties detected by color vision systems. (1999, pág. 64)

⁶¹ Quizás podríamos agregar las *absortancias* (absorptances) a esta lista.

Una consecuencia interesante es que **el realismo antropocentrista *ni siquiera puede recoger la multiplicidad de propiedades detectadas por el sistema de visión humano.***

La definición (c) de Hilbert pretende capturar la idea de que debe haber una propiedad que explique la unicidad de la experiencia de color y, a su vez, la semejanza del orden de los colores (la instanciación de una propiedad de reflectancia X provocará la experiencia de color Y. X e Y covarían mutuamente en un espacio de relaciones cromáticas: si X obtiene un valor determinado provocando la sensación Y, X+1 provocará la sensación Y+1) Por ello, *unicidad y orden de semejanza de los colores* son dos rasgos que las tres definiciones (a)-(c) explican en conjunto. Ya hemos visto la implausibilidad de encontrar una *única propiedad externa* que explique la unicidad de la experiencia de color. Ahora examinaremos si hay algún tipo de *semejanza* entre la experiencia de color y el conjunto heterogéneo de propiedades de las emisiones espectrales que las provocan.

Matthen tomará el argumento de Hardin (Hardin, 1984. Citado en Matthen, pág. 65):

Hardin notes that the resemblance ordering of color as humans experience is an amalgam of a number of similarity and contrast relations amongst the colors that do not (*pace* Hilbert) belong to *any* quality field that exist independently of our perception of color. He concludes that color realism is false: *none* of the things mentioned above has color independently of our perceptions. I do not think that this antirealist conclusion is warranted. However, the facts do, in my opinion, warrant a different conclusion –that color cannot be defined by a resemblance ordering. (1999, pág. 65)

Quizás, entonces, no haya que abandonar el realismo con respecto a la definición del color; tal vez haya que sacar alguna conclusión diferente con respecto al orden de semejanza (fenoménico) de los colores. Hardin establece que hay una diferencia sistemática entre los caracteres de la experiencia visual y las longitudes de onda que le sirven de base. En otras palabras, que existe una discontinuidad entre el aspecto fenoménico de las relaciones de color y el aspecto funcional. Mostrará esta discontinuidad en dos aspectos que nos sirven para relacionar los colores: oponencia y categoricidad:

- **Oponencia:** La estructura de la experiencia ocasionada por los pares oponentes no tiene correspondencia con el ordenamiento de los colores por longitud de

onda en el espectro luminoso, que es independiente del sistema perceptivo. Ejemplos: 1) La similaridad del violeta con el rojo se basa en sus componentes (violeta = rojo + azul). Rojo es cercano al naranja, pero el naranja está en el extremo opuesto del espectro visual del violeta⁶². En realidad, el violeta está más cercano al color verde que al naranja. Sin embargo, el violeta no se ve de ese modo: el proceso oponente distorsiona el orden de los colores tal como está dado por longitud de onda. 2) El color violeta es experimentado como *mezcla* (es binario). Sin embargo, en el espectro luminoso, el violeta ocupa *un* lugar en la longitud de onda: en sí mismo, no es más unitario o más binario que otro. 3) Un mismo color posee variaciones de blanco y negro que son *independientes* de las tonalidades de color. Sin embargo, no percibimos tonalidades de grises *en* los colores. “Este verde es más gris que este otro” no es un juicio que se asevere normalmente acerca de un color: “For these and other reasons, the quality spaces of reflectance and wavelength are simply incommensurable with that of hue, saturation and brightness, which is created by opponent processing” (1999, pág. 66)

- **Categoricidad:** la capacidad de discriminar y dar nombre a los colores varía entre culturas. Los colores, a su vez, tienen connotaciones afectivas: el rojo provoca agitación, mientras que el verde tiende a tranquilizar. Matthen remarca “It makes no sense to say that red would have been agitating or bold in the absence of perceivers” (1999, pág. 67)

Dado que la *categoricidad* y la *oponencia* otorgan a la experiencia de color propiedades que no corresponden a las propiedades que existen independientemente de la percepción (i.e., propiedades del espectro luminoso visible), Hardin concluye que el color es puramente subjetivo. (1999, pág. 67) Matthen, sin embargo, considera que esta conclusión es demasiado fuerte: el argumento de Hardin no permite inferir que el color sea subjetivo; sólo dice que el orden de semejanza de los colores, tal como los experimentamos, no tiene contrapartida con hechos reales. Pero eso no tiene por qué implicar que los hechos que experimentamos *no sean reales* (aun cuando la oposición

⁶² No confundir la “cercanía” o “lejanía” en el espectro visual con la “cercanía” o “lejanía” en el proceso oponente.

entre colores pudiera no serlo): “Why can one hold that experience of red and experience of green track properties that exist independently of us, but that the experienced opposition between them has no real counterpart? Why can we not say that color experience misrepresents the true resemblance relation of colors?” (1999, pág. 68)

La primera conclusión a la que se arriba es que no parece existir semejanza entre el proceso oponente y el color “objetivo”. La segunda conclusión que obtiene Matthen es una lección a partir de la estrategia de Hardin: indicar que el orden de los colores en los pares oponentes no tiene semejanza con el orden del espectro luminoso parece implicar para Hardin que la teoría de pares oponentes resulta *crucial* para definir lo que es un color. Por lo tanto, la falta de correspondencia entre el elemento crucial y la realidad pone en aprietos al realismo, y de allí deriva que el color es puramente subjetivo. Pero, ¿por qué hay que erigir al sistema de pares oponentes como el elemento definitorio del color? La oponencia resulta problemática por dos razones:

The first reason is that opponency is not the only cognitive mechanism that shapes our experience of color. Categoricity is an extremely important feature of human color vision because it is crucial for naming the colors and remembering the colors of objects. (...) Categoricity distorts the hue-brightness-saturation ordering, and the consequence is that the experienced ordering is different from the latter. Hardin’s argument against the realists tells equally against his own contention that that ordering is essential to the colors. (1999, pág. 69)

No hay razones para que la oponencia esté por encima de la categoricidad. Sin embargo, un importante motivo por el cual Hardin (y, aventuramos, quizás, Pautz y Foss) se basa en el proceso oponente es porque *todos los animales que perciben color tienen alguna variedad de ese proceso*. Este motivo es importante y escapa del antropocentrismo. Sin embargo cae en otra variedad de chauvinismo: el *biochauvinismo de la visión*: los colores y su visión sólo pueden definirse a partir de los mecanismos de los seres vivos que ven en color. De modo que una computadora que pudiera generar conceptos de color a partir de una cámara digital *no estaría viendo colores*.

In any case, why should we think that only biological organisms can generate color concepts? If a computer were to generate a set of classifications on the basis of input from a digital “color” camera, it might not come up with a hue-saturation-brightness

structure since the camera does not employ the sampling techniques and differencing characteristic of opponent processing, but encodes colors in a different way. (1999, pág. 69)

¿Puede haber alguna especie extraterrestre que vea en colores? Si así lo fuera, es difícil aceptar que sólo pueden hacerlo en virtud de un proceso de pares oponentes que conserve las mismas relaciones de similaridad entre experiencias cualitativas de color. Robots y vida extraterrestre estarían excluidos de la posibilidad de ver colores. Estas es una consecuencia indeseable si deseamos definir la experiencia del color de una manera general. Matthen deriva una cuarta conclusión a partir del análisis de los argumentos de Hardin:

Conclusion 4: Anthropocentric definitions that define color in terms of the *resemblance ordering* or other property structure found in human experience are defeated by the multiplicity of possible color resemblance orderings. (1999, pág. 69)

En resumen, las cuatro conclusiones advierten acerca de los peligros de definiciones en color tanto en términos explícitamente antropocentristas (conclusión 1), como aquellas que lo hacen en términos de un antropocentrismo flexible que parte del caso humano para hacerlo extensivo a otras especies (conclusión 2), como aquellas que buscan una propiedad real unitaria para definir el color que a la postre es difícilmente compatible con las propiedades que detectan otras especies -e incluso el hombre- (conclusión 3), como aquellas que se basan en el proceso oponente (conclusión 4).

Estas cuatro conclusiones nos invitan a entender el color a partir del *realismo pluralístico*: abandonar la búsqueda de semejanzas *reales* entre un sistema perceptivo determinado (por ejemplo, el humano) y un conjunto de propiedades unitarias presentes en la naturaleza. La visión del color puede detectar un gran número de propiedades (no sólo reflectancia en las superficies); no hay una única estructura de detección posible ni hay un único modo de darse la experiencia de color. Podría preguntarse, en consecuencia, si existe algún factor convergente que pueda unificar el concepto de *color*: “Pluralistic realism also needs to be clarified with respect to the question of definition that we have been considering: what makes all of these different properties color?” (1999, pág. 72)

Existe una versión del realismo para la cual tal convergencia está dada por las cualidades ambientales que son la base para todo tipo de visión en color. Las cualidades están “ahí afuera”, y la visión en color las “captura” (*catch*). “If it is so, catching colors must constitute the essence of color vision. If there is a fixed number of colors ‘in the world,’ variation is limited: catch them all and you have full color vision, catch only some of them and you are color-deficient” (1999, pág. 73). En todo caso, ¿es necesario, para el pluralismo realista, aceptar que hay un número limitado de colores y que los organismos simplemente los “capturan”?

Una manera de responder a la pregunta anterior es analizar el rol que cumplen los colores para un determinado organismo. ¿Qué tipo de ventaja representa la visión en color, a diferencia de especies que no la poseen? Matthen encontrará que ninguna de las ventajas usualmente postuladas parece decisiva: en situaciones para las cuales la discriminación de color es relevante, el entorno suele proveer otros modos de segmentación de una escena. El color no se necesita para construir los límites de los objetos, ni para generar contraste entre ellos (1999, pág. 74). El color podría servir para distinguir elementos cuando el fondo no ofrece un gran contraste, o para reconocer objetos en lugar de diferenciarlos. Ninguna de estas ventajas, juzga Matthen, parecen decisivas.

These conclusions challenge the idea that color vision evolved because there are determinate qualities out there, *the* colors, so universally salient that widely divergent species converge on them. (...) It is more appropriate to think that color vision is advantageous because it enhances visual discrimination abilities already in place. For example, it helps humans extends their ability to segment a scene to situations where the background is dappled and brindled, it helps birds find their way when there is heavy cloud cover obscuring the direction of the sun, and it helps bees find the part of the flower with pollen. These different kinds of organism have quite different visual discrimination abilities, and use them for different tasks. Why then should we expect that they will all be helped by “catching” some elements out of a fixed store of colors? (1999, págs. 75-6)

Cada organismo modela un específico concepto de color, adaptado a las características ambientales específicas en las que está inserto. Ese concepto de color no es, sin embargo, una “clase natural” que se “captura”. A este respecto es instructiva la advertencia de Daniel Dennett:

Los colores no son «clases naturales» precisamente porque son un producto de una evolución biológica, la cual tiene una tolerancia con los límites difusos cuando produce categorías que horrorizarían a cualquier filósofo amante de las definiciones buenas y claras. Si la vida de alguna criatura dependiera de reunir la luna, el queso azul y las bicicletas en una sola categoría, pueden estar seguros de que la Madre Naturaleza encontraría el modo de que esa criatura «viera» todas estas cosas «intuitivamente como la misma clase de cosas» (1995, pág. 393n)

Según Dennett y Matthen, no se trata de “capturar” características externas, sino de adaptar un sistema que se aproveche de ciertas propiedades para adaptarlas a las necesidades del organismo:

Some organisms have a colored oil-drop in front of their retinas, in order to enhance hue discrimination in specialized lighting situations. Should we say that they are distorting the true colors out there, like a person who dons cerulean spectacles? (1999, pág. 76)

Un organismo cuya retina posee una gota de aceite coloreada, ¿tiene una percepción *distorsionada* de los *verdaderos* colores? No parece plausible. Por ello, Matthen propone una definición de color caracterizada no a partir de los estados cualitativos de la visión (los qualia de los colores), sino a partir de los *input* visuales y su funcionalidad en un ambiente determinado. ¿Qué es la visión del color?:

An organism possesses color vision if it uses wavelength discrimination as an indispensable part of its construction of visual representations of environmental features. (1999, pág. 78)

¿Qué es un *quale* visual?

A color quality is one that is generated from the processing of wavelength-sensitive data and becomes available only as a result of such processing. (1999, pág. 78)

Estas definiciones satisfacen los siguientes requisitos:

- a) No están constreñidas a situaciones puramente biológicas. El “organismo” al que se alude en la primera parte de la definición puede ser un “organismo artificial”.
- b) No se presupone una estructura oponente para generar la cualidad del color.

- c) El quale visual no está definido a partir de una determinada especie, sino a partir del procesamiento de datos acerca de la longitud de onda de un espectro luminoso.
- d) La fenomenología visual depende de las especializaciones adaptativas de cada organismo, y no puede cruzar barreras filogenéticas: los colores que ve una abeja no son conmensurables con los que ve un humano, dado que entre ambos no hay un ancestro común con visión en color⁶³.

Matthen reconoce, sin embargo, que las definiciones dadas no nos ayudan a elaborar una concepción clara de “cómo es” (what is it like to be) ver un color. “When we have nothing but color to go on, we have no basis for projecting what color categories might be like in an organism quite different from ourselves” (1999, pág. 82) Las definiciones pretenden delimitar el ámbito de la visión de color, pero esto no implica que haya una unidad en la experiencia de color. A nuestro juicio, este resultado parece un poco sorprendente. Si no hay una unidad entre diferentes especies en lo que respecta a la experiencia de color, podemos suponer que la *experiencia de amarillo* (independientemente de la longitud de onda luminosa) podrá ser dramáticamente diferente entre especies. La específica cualidad del amarillo sólo puede valer para una especie: la humana, pues no podemos elaborar un concepto que nos permita decidir si para las palomas o las abejas ver amarillo es equivalente a una experiencia del mismo tipo que las experiencias humanas al ver amarillo. ¿Es el amarillo que ven las palomas más parecido a nuestro naranja, o a nuestro violeta? ¿O quizás a ninguno de los dos? Nuestros conceptos de color parecen dependientes de la categoricidad humana y no son claramente proyectables entre especies. No sólo resulta imposible pensar (imaginar) cómo es ser murciélago; es incluso difícil imaginar cómo es ver amarillo –el mismo amarillo que estoy viendo ahora- si fuera una paloma que está teniendo una experiencia de amarillo.

⁶³ “...[I]t is a mistake to think that there must be a characteristic kind of datum that is shared by [different] organisms. Color phenomenology may arise out of specialized adaptations that do not cross phylogenetic boundaries. Because humans and honeybees have no common ancestors with color vision, there is no reason to think that the two color vision mechanisms developed as a response to comparable environmental challenges” (Matthen, 1999, pág. 80-1)

2.9. Acerca del parecer: ¿Adónde hemos llegado?

Hemos partido de un análisis acerca de la noción de *parecer* tal como está presente en afirmaciones del tipo “me parece que”, “se ve como”, “lo veo (de tal o cual manera)”, “cómo es ser...” y términos en inglés afines (“it seems to me”, “it looks...”, “what is it like to be”) para desembocar en un análisis del *parecer* en la visión del color. Este recorrido –cuyas consecuencias apenas hemos sugerido– tiene algunos mojones que sirven de lazo entre los temas y autores reseñados: partimos de la pesadilla epistemológica de Raymond Smullyan, en la cual una persona *no podía decir con verdad* que creía lo que afirmaba que creía. El diálogo entre Dennett y Otto reafirmaba el horror de esta pesadilla: Otto no podía hablar con sentido acerca del anillo que le *parecía* que había, y reclamaba una explicación que Dennett sistemáticamente le mostraba como innecesaria y sin sentido, dado que Otto pretendía hablar de un estado cualitativo que no tenía lugar y que minaba su juicio acerca de su parecer. Dennett, en *Quining Qualia* identifica el *modo en que las cosas nos parecen* con los *qualia*; Gibbons busca razones para separar ambos conceptos, y las conclusiones (cautelosas) de su examen parecen argumentar en contra de los *qualia*: todo lo que hay es un *parecer*₁, dirigido hacia propiedades externas. Ese *parecer*₁, según Jeff Foss, puede explicarse sin ningún resto a partir de la teoría de pares oponentes. Una explicación tal es, en rigor, todo lo que se requiere, y de hecho la explicación ya ha sido dada. *Lo que nos parece* un color no necesita de una extraña esencia cualitativa primitiva no superveniente. Adam Pautz afirmará que el sistema de pares oponentes indica que la experiencia (lo que nos parece₁) el color depende directamente de los pares oponentes cuyos correlatos neurales se encuentran en el cerebro: el color, a diferencia de Foss, es una cualidad mental primitiva para la cual no es posible encontrar semejanza en el mundo externo. Byrne y Tye, por su parte, replicarán que no se han establecido los correlatos neurales del proceso oponente, que no se entiende en qué sentido lo que ocurre en el cerebro determina la visión del color (dado que Pautz no niega la influencia del mundo exterior); luego Mollon destacará que la teoría del proceso oponente fue postulada por Hering a partir de observaciones puramente fenomenológicas y que no existe evidencia de que tal fenomenología se

vea reflejada en algún proceso funcional posreceptor. Finalmente, con Mohan Matthen vimos que las definiciones usuales de ‘visión de color’ contienen un componente antropocentrista (entre los cuales se encuentra el proceso oponente) que no permite capturar la experiencia del color en seres de otras especies, o incluso de otros planetas u otra manufactura, como las computadoras.

No hemos extraído las consecuencias de este recorrido, aunque ahora comenzaremos a hacerlo. Retomemos la definición de Mohan Matthen sobre el *quale* visual:

A color quality is one that is generated from the processing of wavelength-sensitive data and becomes available only as a result of such processing. (1999, pág. 78)

Hemos examinado las ventajas de esta definición, pero a nuestro juicio posee una desventaja evidente.

Supongamos que se descubre una especie que pueda procesar datos relativos a la longitud de onda luminosa, pero para la cual la experiencia del color fuera similar a nuestra experiencia de un *sonido*. Allí donde nosotros *vemos rojo*, los individuos de esta especie *oyen agudo*. Quizás no haga falta explorar casos hipotéticos: el curioso fenómeno de la *sinestesia* nos provee de un ejemplo al respecto. Los estados cualitativos *parecen*₁ de una modalidad diferente a la que de hecho son: en la sinestesia se pueden ver olores, oír colores y ver sonidos. Podemos suponer que un día, por un curioso fenómeno evolutivo, todos los seres humanos nacemos con dos modalidades invertidas: vemos los sonidos y oímos los colores. Esto significa que, en el caso de oír los colores, se ha generado un estado cualitativo a partir del procesamiento de la longitud de onda que ha impactado en los sentidos. Se cumplirían perfectamente las condiciones de la definición de Mohan Matthen. Pero, ¿se puede llamar a esa experiencia una *experiencia de color*? Parece que no o que, en todo caso, el procesamiento de la onda luminosa no es suficiente para caracterizar a la cualidad del color, *aun* entre especies diferentes. Matthen advierte que no pretende capturar la fenomenología de la experiencia de cada especie en particular, pero la posibilidad que proponemos podría hacerse extensiva incluso a todas las especies. ¿Podríamos decir que alguien ve en colores incluso en ausencia de todas las características fenomenológicas típicas de la visión en color? Nos parece que la definición de Matthen

es suficientemente amplia como para escapar del antropocentrismo e incluso del biochauvinismo, aunque no es lo suficientemente estrecha como para excluir inversiones en la fenomenología intermodal. Debería establecerse alguna restricción para evitar que fenómenos de otras modalidades puedan surgir como consecuencia de la actividad de una modalidad sensorial diferente. Peter Ross (2001) propone un conjunto de condiciones necesarias para distinguir un sentido de otro. Estas condiciones pueden resumirse en tres puntos:

- a) Un sentido se especializa en la percepción de un rango de propiedades de objetos físicos (por ejemplo, color o sonido) (Condición de la propiedad)
- b) El sentido es receptivo a una clase distintiva de estímulos. (Condición del estímulo)
- c) El sentido está asociado con el trabajo de una clase de órganos. (Condición del órgano sensorial) (2001, pág. 500)⁶⁴

La satisfacción de las tres condiciones es lo que garantizaría la unidad de un sentido. Una diferencia en las condiciones indica que la experiencia proviene de (al menos) dos sentidos diferentes (si dos experiencias sensoriales S_1 y S_2 comparten a y b, pero no c, entonces S_1 y S_2 pertenecen a sentidos diferentes). Sin embargo, la primera de estas condiciones – como el mismo Ross admite- puede ponerse en peligro a partir de los ejemplos de sinestesia:

(...) [A] rare condition called synaesthesia suggests an interaction among senses which is so extreme as to undermine use of the (...) property condition. (...) Little is currently known about synaesthesia. But it is sometimes described as involving, for example, seeing sounds or hearing colours. If such perceptual experiences occur, senses do not specialize in perceiving distinctive range of properties. The properties supposedly distinctive of each of the senses can be directly perceived by other senses. (...) I maintain that our current murky understanding of synaesthesia does not provide any adequate basis for rejecting use of the (...) property condition in distinguishing senses. (2001, pág. 503)

La sorprendente respuesta de Ross consiste en dejar de lado el fenómeno de la sinestesia sobre la base de nuestro desconocimiento al respecto, pero es importante destacar que, si de hecho la sinestesia es tal que, literalmente, se puedan escuchar

⁶⁴ Modificamos ligeramente el original, dado que en el texto de Ross se está dando respuesta a un experimento mental que hemos omitido.

colores y ver sonidos, constituiría un contraejemplo de las condiciones a-c para distinguir un sentido de otro. Más importante aún: las condiciones establecidas no pueden dar cuenta de las inversiones en la fenomenología intermodal. En ese sentido, no agregan una restricción ulterior a la definición de Matthen.

Aun cuando, siguiendo a Ross, dejemos de lado el fenómeno de la sinestesia en casos humanos, por tratarse de un fenómeno cuyo entendimiento es escaso y del cual no se pueden sacar conclusiones acerca de la distinción entre sentidos, existe al menos un caso en el que la sinestesia parcial parece estar apoyada por datos empíricos. Horst (2007, pág. 161) expone el caso de los hurones, en comparación con el de los humanos:

In ferrets, as in humans, auditory signals pass through the thalamus before reaching the auditory cortex. But whereas these connections are present at birth in humans, they develop after birth in ferrets, and hence interventions are possible in neonatal ferrets. Sur found that if he cut the connections of the auditory stream to the thalamus, the optic nerve would grow connections to both “visual” and “auditory” areas of the cortex (Sur, Garraghty, and Roe 1988). More surprising still, in these ferrets the auditory cortex would develop the “pinwheel” organization of cells normally found in the visual, but not the auditory, cortex. These formations were not as numerous or as orderly as those in the visual cortex, but were nonetheless distinctive anomalies not found in normal auditory areas (Sharma, Angelucci, and Sur 2000). The portions of the ferrets’ brains that would normally have served as an auditory cortex instead became (functionally) a second visual area and developed a structure with features associated with a normal visual cortex. Sur’s conclusion is that such brain regions exhibit a significant degree of developmental plasticity and acquire their “function” only through development and experience”

Si mediante una intervención es posible cortar la unión del canal auditivo con el tálamo, el nervio óptico desarrolla células tanto en el área visual como en el área auditiva de la corteza cerebral, e incluso se desarrolla en el área auditiva una organización celular similar a la del área visual. Este hecho probaría la plasticidad funcional del cerebro para “hacerse cargo” de capacidades que no están disponibles, y en gran medida dejan fuera de juego las condiciones a) - c) que Ross exponía para distinguir los sentidos. Se necesita, entonces, un componente adicional para que tal distinción sea efectivamente una distinción sensorial: quizás ese componente faltante sea el *específico estado cualitativo –fenomenológico- de una experiencia de color*.

Si la noción de ‘visión de color’ no puede dar cuenta de la típica característica de la *experiencia* de color (porque podría confundirla con la experiencia de un sonido, por ejemplo), entonces está faltando algo importante. En todo caso, el procesamiento de datos de longitud de onda parece ser una condición necesaria, aunque no suficiente, de la visión. Pero, ¿es realmente una condición necesaria?

Ned Block (Mental Paint, 2003) ha llamado la atención acerca de la experiencia de fosfenos⁶⁵. Cuando se presiona el ojo cerrado, se tiene una sensación visual de colores danzantes sin que exista luz presente. Se obtiene una experiencia fenomenológica de visión de colores sin que se cumpla la condición establecida por Matthen: procesamiento de datos de longitud de onda. Si esta interpretación es correcta, hay al menos *un* caso que escapa de la definición de Matthen. La experiencia de fosfenos es un fenómeno acotado y bastante marginal, pero si suponemos la existencia de una especie cuyos ojos no tienen contacto con la luz (porque, por ejemplo, por razones adaptativas se vio confinada a una vida en cavernas oscuras y sin acceso a ningún tipo de iluminación artificial), cabría la posibilidad de que sus experiencias visuales paradigmáticas fueran la visión de fosfenos. Los miembros de esa especie podrían hablar acerca de la riqueza de colores que ven e incluso ponerles nombres. Su especie podría generar una muy fina conceptualización acerca de los colores fosfénicos. Supongamos que algunos miembros de esta especie llaman “*Zahu*” al celeste-fosfeno. Si se lleva a uno de ellos a la luz del día y se le muestra el cielo, es probable que encuentre identidad entre el celeste del cielo y el *zahu* del fosfeno. Podría decir algo como esto: “¡Es *el mismo color* *zahu* que veo en la caverna!”. En este punto, el partidario de la definición de Matthen debería advertirle algo. Si tomamos literalmente la definición de Matthen, los “colores” vistos a través de fosfenos *no serían cualidades de color*, pues no están originados por el procesamiento de datos de longitud de onda. Por lo tanto, el celeste que está viendo en el cielo no puede ser idéntico al *zahu*. Deberíamos aclararle que la riqueza conceptual aprendida en la caverna, la fina trama de nombres de colores-fosfeno, no tiene correspondencia con esos colores que está viendo ahora y que debería mantener separadas ambas familias de conceptos... ¡Aun cuando todos los individuos la especie encontraran diferencias y similitudes

⁶⁵ Mencionamos brevemente en qué consiste esta experiencia en el apartado 2.5

cualitativas que son en todo idénticas entre las experiencias-fosfeno y las experiencias en las que interviene la luz! Por ejemplo, los individuos de esta especie encontrarán que el celeste-fosfeno es más parecido al azul-fosfeno que al rojo-fosfeno, similaridad que todos apreciarán como idéntica en las experiencias en las que interviene la luz. Pero a pesar de estas curiosas similitudes, si sigue la definición de Matthen debería desistir de seguir llamando “color” a lo que veía mientras estaba en la caverna.

¿Cómo puede lidiar la definición de Matthen con ilusiones ópticas como las del anillo rosa de Otto, o el círculo verde? En ambos casos, *no* hay una determinada longitud de onda luminosa que provoque *esa* experiencia. No hay correspondencia entre el verde visto y el procesamiento de una frecuencia luminosa de, por ejemplo, 525nm. El verde visto en la ilusión óptica no calificaría como cualidad de color. Dado que Matthen establece su definición a partir del *input* luminoso, ninguna experiencia semejante a la visual que no cumpla con ese *input* podría llevar el título de experiencia visual. En estos casos, quizás, nos *parece* que vemos colores, pero en rigor no los vemos. Después de un largo rodeo hemos vuelto a instaurar el argumento de Dennett contra Otto.

Todos los verbos de percepción (ver, oír, oler, gustar, sentir) implican en principio el éxito perceptivo. Si digo que veo una superficie roja, estoy emitiendo un juicio cuyo contenido es, parcialmente, que *hay* una superficie roja. Si tengo razones para suponer que esa superficie roja que veo en realidad no existe (porque estoy alucinando o porque sospecho que se trata de una ilusión óptica), el verbo “ver” en este caso se torna ambiguo. Podría ser cauteloso y decir “me parece que veo una superficie roja”, pero esto no sería del todo exacto si a pesar de mi cautela sé que estoy teniendo la experiencia fenomenológica del color rojo (No *me parece* que estoy teniendo la experiencia de rojo; simplemente la estoy teniendo). Quiero ser cauteloso con respecto al éxito perceptivo, pero no tiene sentido aplicar esta cautela a mi propia experiencia fenomenológica, pues yo no tengo dudas de que estoy teniendo esa experiencia: de lo que tengo dudas es de que esté viendo –es decir, percibiendo– una superficie roja. ¿Tenemos derecho a seguir utilizando el verbo “ver” cuando sabemos que el objeto visto no existe, ya sea porque se trata de una alucinación, una ilusión óptica o un fosfeno? Si tomamos la definición de Matthen, parece que no. Las

experiencias de “ver color” vienen implicadas por el éxito perceptivo; esto es, por el procesamiento adecuado de los datos de longitud de onda. En ausencia de tales datos, no puede haber éxito perceptivo; todos los juicios elaborados se quedan sin sustento y son, por lo tanto, falsos. “Estoy viendo rojo” es falso si no está presente, como elemento causal, el dato de una longitud de onda luminosa entre 620 – 750nm. ¿Estaría mintiendo una persona que dijera que está viendo rojo en ausencia de un *input* luminoso? No, pero su juicio sería erróneo *aun cuando de hecho estuviera teniendo una experiencia típica de rojo*. Todavía podría decir que *le parecía*₁ que había una superficie roja. Recordemos que Dennett le prohíbe a Otto decir, siquiera, que *le parece*₁ que hay un anillo rosa. Si se extrema el ejemplo con la restricción de Dennett, llegamos a un resultado inaceptable:

- Si, ante una experiencia alucinatoria, ilusoria o fosfénica, afirmamos que *vemos* un determinado color, tal afirmación es falsa. (Pues a la visión le falta el elemento causal, el *input* correspondiente)
- Si corregimos la afirmación anterior y decimos que *nos pareció* que habíamos visto un color, entonces caemos en el sinsentido (Si aceptamos las reglas de Dennett al respecto: no se puede hablar de la fenomenología cuya causa es algo que no existe, como el anillo rosa o el miembro fantasma)

De este modo, parece que o bien se inventa un lenguaje diferente que no incluya verbos de parecer ni de percibir, ni adjetivos de color (el rojo de un fosfeno debería llevar otro nombre, pues en rigor no es un color), o bien se abandona todo intento de hablar de las experiencias visuales si no hay un correlato causal correspondiente.

Pero, ¿de qué están hechas nuestras experiencias visuales? ¿Cómo se pueden caracterizar? Gibbons nos decía que toda nuestra experiencia está dirigida hacia afuera (pág. 172):

Where does the closest red thing seems to be? If you are not visible bleeding, and you are like the rest of us, it will seem to you as though the closest red thing is at some distance from you. It seems to be outside of you. It does not seem to be inside of you. Of course, it does not follow from this that it really is outside of you. (...)

Where is the redness seems to be? “There!” you say, pointing to the external world. Of course, it might not really be there. That is just where it seems to be. (Gibbons, 2005, pág. 413)

¿Es posible señalar un “allí afuera” cuando hablamos de fosfenos? La experiencia fosfénica no puede señalarse en un espacio exterior. En ese caso, el “color rojo más cercano” que yo pudiera ver quizás no sea uno que esté a una cierta distancia de mí, como pretende Gibbons, sino *en mí*. Como señala Block, “the diaphanousness of perception is much less pronounced in a number of visual phenomena, notably phosphene – experiences” (2003, pág. 13). Cuando se tiene una experiencia visual, es posible atender a la propia experiencia sin referencia al hecho externo que la origina. Esto es notorio en el caso de los fosfenos:

Can you attend to those sensations? I believe I can. Even if you can't attend to them, are you aware of them? According to the representationist, all awareness of those sensations could consist in is awareness of the colored moving expanses that are represented by them. My view is that one can be aware of something more. (2003, pág. 13)

Cuando Dennett negaba que hubiera fenomenología, estaba diciendo que el contenido de la sensación es puramente intencional; que no hay un aspecto intrínseco – fenoménico- en ese contenido. Block, en cambio, entenderá que hay una diferencia entre el contenido intencional y el propio fenómeno de la sensación:

I think that (...) at least sometimes when we introspect visual experience, we are attending to the phenomenal properties of experience, not the fact that they have certain intentional properties. (2003, pág. 14)

El propio fenómeno es, según Block, atendible en sí mismo. Nótese, sin embargo, que la definición de “visión en colores” de Matthen dejaba afuera el carácter fenoménico del color. Recuérdese, también, que la explicación del *quale* –esto es, del fenómeno o del *way it seems*- de la visión del color *amarillo* quedaba suficientemente explicada, según Foss, a través de la teoría del proceso oponente. Pero, por un lado, el proceso oponente no puede hacerse extensivo a una definición general de “visión de color”, sino a una puramente humana y, por otro lado, tal proceso parece ser una caracterización fenoménica. O bien, desde un lado, se elimina al fenómeno del color de la definición de “visión de color”, o, desde otro lado, se explica la visión del color *desde* el propio fenómeno (el proceso oponente), pero sin que se puedan encontrar los correlatos funcionales de dicho fenómeno. Tanto Pautz como Byrne-Tye y Mollon han advertido que la teoría del proceso oponente está lejos de encontrar una base

funcional. Por otra parte, mecanismos funcionales diferentes dan lugar a fenomenologías diferentes, pero, tal como Matthen nos previene, la fenomenología inter-especies no es comparable: “When we have nothing but color to go on, we have no basis for projecting what color categories might be like in an organism quite different from ourselves” (Matthen, 1999, pág. 82); esto es, si partimos del fenómeno, de *lo que nos parece*₁ el amarillo, es imposible determinar si a otra especie (por ejemplo, la paloma) ve *el mismo amarillo* que vemos nosotros *cuando tenemos el parecer*₁ de amarillo. Así, los nombres de los colores *en tanto los experimentamos*, sólo tienen validez humana y difícilmente puedan ser proyectables a otras especies que también ven color. Pero quizás ni siquiera la *experiencia* ni la *fisiología de la percepción* de un determinado color puedan mantenerse como una base unívoca para hablar del color. Matthen concluye que el color es una cualidad *no unívoca*, ni en sus bases fisiológicas ni en su fenomenología:

Color is a disunity, then. There is neither a phenomenology of color vision nor a set of shared concepts that defines color. There is a commonality in the informational material from which color concepts are constructed, and a consequent functional commonality in the mechanisms that are needed to gather this information, but no mind-independent property that all color perceivers track or detect, no one ecological problem they all try to solve. Color is a heterogeneous collection of perceptual concepts generated from wavelength-sensitive data for a variety of purposes by cognitive systems with different neurocomputational structures and evolutionary histories (1999, pág. 84)

¿Cómo definimos, pues al color? No contamos con bases para dar una definición general de la fenomenología del color ni de los procesos funcionales ni de las propiedades que detecta.

En un sentido, tampoco podemos hablar del color como una propiedad unívoca ni siquiera en el ámbito humano. Quizás para usted, que ve en condiciones óptimas y no tiene defectos en la vista, el “verde sin mezcla” sea diferente del mío, en las mismas condiciones perceptivas. Block remarca que existen diferencias físicas en la propia estructura retiniana que provocan diferente fenomenología según el sexo, la raza o la edad:

There are three kinds of cone in the retina that respond to long, medium and short wave light. (...) Two normal people chosen at random will differ half the time in peak

cone sensitivity by 1-2 nm (nanometers) or more. (...) This is a considerable difference, given that the long wave and middle wave cones only differ in peak sensitivities by about 25 nm. Further, there are a number of specific genetic divisions in peak sensitivities in the population that are analogous to differences in blood types (in that they are genetic polymorphisms, discontinuous genetic differences coding for different types of normal individuals). The most dramatic of these is a 51.5%/48.5% split in the population of two types of long wave cones that differ by 5-7 nm, roughly 24% of the difference between the peak sensitivities of long and middle wave cones. (Neitz and Neitz, 1998) This characteristic is sex-linked. The distribution just mentioned is for men. Women have smaller numbers in the two extreme categories and a much larger number in between. (2003, págs. 24-5)

Estas diferencias son *funcionales*, (están relacionadas con factores genéticos que afectan de modo diferenciado la respuesta sensitiva de los conos). También hay diferencias *fenoménicas*. Block (2003, pág. 26) menciona un estudio clásico de Jurvich, Jameson y Cohen (1968) en el que un grupo de cincuenta sujetos expone la localización (fenomenológica, i.e., de acuerdo a cómo a cada sujeto "le parece" dicha localización en el espectro) del verde puro. La tabla de localizaciones varía considerablemente (Figura 11):

5 sujetos localizaron el verde unitario a 490 nm	
11	500nm
15	503nm
12	507nm
5	513nm
2	517nm

Figura 11. (Extraído de Block, 2003, pág. 26)

La tabla muestra un enorme rango de variación entre la fenomenología y el correspondiente espectro luminoso. Quien negara la fenomenología debe decir, dado que el contenido intencional es la *única* nota distintiva del carácter fenoménico, que a pesar de todas estas diferencias el *contenido representacional* no variará. Pero:

The upshot is that there is an empirically based argument for a conclusion—what one might call "shifted spectra"-- that, while not as dramatic as an inverted spectrum, has much the same consequences for representationism and for the issue of whether

there are uniform phenomenal characters corresponding to colors. There probably are small phenomenal differences among normal people that don't track the colors that are represented. Genders, races and ages probably differ by shifted spectra. Thus, if representationism is right, if aquamarine things look aquamarine to men, they probably don't look aquamarine to women. And if aquamarine things look aquamarine to one race or age group, they probably don't look aquamarine to others. In sum: If representationism is right, color experience probably cannot be veridical for both men and women, both blacks and whites, both young and old. Hence representationism is not right. (2003, pág. 27)

Block llama "representacionalista" a quien mantiene que el contenido intencional de una experiencia es *todo* lo que se requiere para explicarla. Quien mantiene esta posición, no podrá explicar casos en los que hay sutiles diferencias en la fenomenología que no pueden codificarse como diferencias en la representación o contenido intencional. O, en todo caso, deberá admitir que el contenido intencional varía de acuerdo al sexo, la edad, la raza o incluso características puramente idiosincráticas, aun cuando la manifestación de dicho contenido intencional (en un juicio) pudiera no mostrar diferencias.

También, como vimos inmediatamente más arriba, hay diferencias en la propia estructura sensitiva de los conos. Si la afirmación de Matthen es cierta, "When we have nothing but color to go on, we have no basis for projecting what color categories might be like in an organism quite different from ourselves" (1999, pág. 82) (Y también: "A lot of our intuitions count for little in the unfamiliar territory of inter-species comparisons" [pág. 61]), el panorama es aun más *disunívoco*, porque incluso dentro de nuestra propia especie, en vista de las diferencias por sexo, edad, raza y más aun, diferencias idióticas surgidas por la categorización que combina el *parecer*₁ con el *parecer*₂, un color parece confinado a una experiencia peligrosamente particular, indefinible e incomunicable. *La experiencia de color sólo parece definirse de acuerdo a cómo la experimenta un individuo X en el tiempo Y* lo que nos deja con una afirmación tautológica: *una experiencia de color Y para un individuo X se define por lo que experimenta como color un individuo X en la situación Y*. La sensación de color, en vista de su heterogeneidad, «disunidad», variabilidad e inasibilidad no parece una criatura conceptual más manejable que la fenomenología (y el polémico *quale*) que se pretende desterrar desde la estrategia dennettizadora. Con Matthen y Dennett, vimos

que tampoco hay unidad con respecto a lo que debe considerarse como base para definir qué es un “color”, pues es necesario entenderlo en función de las necesidades de una especie. Por otra parte, la propia caracterización fenoménica del color parece ser sumamente idiótica, restringida a lo que un individuo de la especie X experimenta aquí y ahora. La posibilidad de reducir esta experiencia idiótica, sumado al resultado de Matthen de que no es posible comparar dos estados fenoménicos entre especies, nos deja con campos disunificados que difícilmente puedan reducirse uno al otro: por un lado, las experiencias de la visión de color parecen diferentes en cada especie, en cada individuo e incluso en diferentes etapas de la vida de cada individuo. Por otro lado, no hay un modelo funcional satisfactorio para entender en qué consiste el color sin referencia a la experiencia de ciertos individuos con respecto ese color. Quizás todo lo que tengamos es una miríada de entidades fenoménicas disunificadas, y una miríada de hechos externos que sean las causas no unívocas de ciertas experiencias fenoménicas. Resulta difícil operar una reducción entre este conglomerado disunificado. Tres siglos después del fracaso de la reducción del programa mecanicista clásico, nos encontramos con problemas similares o quizás peores: la eliminación del componente fenoménico conlleva la reaparición de dicho componente; la reducción de la fenomenología a su estructura funcional no se puede llevar a cabo mientras el campo al que se pretende reducir (la teoría del proceso opoente) se mantenga también dentro de la fenomenología. Las causas externas que provocan una experiencia de color no son únicas; la propia fenomenología difiere entre especies y entre individuos; las causas externas y la fenomenología no poseen relaciones de semejanza. Tal vez es hora de aceptar que los intentos de reducción y de unificación son imposibles para la elaboración de la explicación de la naturaleza de la conciencia.

2.10 Modelos explicativos disunificados.

The macroscopic objects we attempt to treat in classical mechanics are enormously complicated in both structure and behavior. Any practical vocabulary must be strategically framed with these limitations firmly in view. To be able to discuss such assemblies with any specificity, our stock of descriptive variables must be radically reduced, from trillions of degrees of freedom down to two or three. – Mark Wilson, Wandering Significance, (pág. 184)

Si hacemos un balance del recorrido que hemos hecho en esta segunda parte, parece que la estrategia dennettizadora se ha encontrado con un camino sin salida: los mecanismos funcionales que se proponen como candidatos para las explicaciones conceptualmente adecuadas (CAEs) de los fenómenos de la experiencia cualitativa del color se muestran imposibilitados de otorgar una caracterización transparente del aspecto fenoménico en términos de estados funcionales. Uno de los candidatos más recurridos, el proceso oponente, consiste en una clasificación puramente fenoménica, cuya semejanza con estados funcionales de la estructura perceptiva y/o propiedades de la luz no puede siquiera establecerse. No se puede explicar la específica fenomenología del color. La estrategia reductiva no se ha llevado a cabo en absoluto. Esta imposibilidad de reducción puede hacerse extensiva a otros aspectos fenomenológicos y no sólo al que hemos tomado como ejemplo (el de la visión de colores). Horst (2007) diagnostica que

Localization studies, marshaling evidence from trauma cases, brain imaging, single-neuron sampling, and animal experiments, reduction and supervenience do lead to conclusions about what parts of the brain are implicated in some special way in the exercise of various psychological capabilities. But this does nothing to explain the fact that their related psychological states have a phenomenology at all; nor does it decide the question of the metaphysical status of these special connections for example, whether they are instances of identity, supervenience, or mindbrain causation. Psychophysical laws like the Weber laws express regularities in the relations between stimuli and percepts, but do not explain why the percept has the precise qualitative character it has, nor indeed why it has any qualitative character at all. Models of early color vision explain things like metamers, the three-color process postulated by Helmholtz, color opponency, and the shape of the human color solid. But again, they do not explain why any of this involves a phenomenology, or the specific phenomenology it (2007, pág. 38)

La fenomenología aparece inexplicada, i.e., irreducible en general y no sólo en los casos de la visión del color. No se trata, al parecer, de una característica local o puntual, sino de una característica global: la conciencia en sí misma se muestra como irreductible.

Podría pensarse que este diagnóstico abonaría la posición nagelizadora: la conciencia posee un *status* especial y es el *único* caso en que no se puede operar la reducción. Sin embargo, quizás tampoco esta posición resulta adecuada.

Steven Horst (2007) analiza con profundidad el problema de la reductividad en la ciencia natural. Los programas naturalistas intentan “naturalizar” la conciencia operando una reducción tal como se ha llevado a cabo en otros casos históricamente exitosos. Para cumplir con este intento, se necesita presuponer, en primer lugar, que la mente es *natural*. Sin embargo no es claro qué debe entenderse por “natural”: en principio, se supone que un dominio D es natural si todas las características de D pueden acomodarse en el marco de la naturaleza tal como es entendida en el ámbito de las ciencias naturales. Esta idea de la *acomodación* ofrece, sin embargo, tres ejes de ambigüedad que tipifican tres tipos de naturalismo. Los ejes deben dar respuesta a tres tipos de problemas:

- I. Whether the “accomodation” in question is a sort of explanation or a type of metaphysical determination.
- II. How we are to understand “the framework of nature as it is understood by the natural sciences”, and
- III. Whether the general schema⁶⁶ is understood as a positive claim (that the mind can be so accomodated) or as a normative claim (that it must be so accomodated, or else some dire consequences follow) (2007, pág. 14)

No es fácil determinar qué clase de naturalismo afirman autores que dicen ser “naturalistas”, dado que puede haber una gran variabilidad *en un mismo autor* con respecto a la respuesta que da a los puntos I-III. Por ejemplo, con respecto al punto II no queda claro cuál es la respuesta de Dennett acerca del rol de las ciencias naturales. Al final de *La Conciencia Explicada*, Dennett describe lo que ha hecho en todo el libro:

No puede decirse que haya sustituido una teoría metafórica por una teoría *no* metafórica («literal», «científica»). La verdad es que todo lo que yo he hecho no es más que sustituir una familia de metáforas e imágenes por otra, cambiando el teatro, el testigo, el Significador Central, el figmento, por el *software*, las máquinas virtuales, las Versiones Múltiples, el pandemónium de homúnculos. Así que no es más que una guerra de metáforas, me dirán ustedes, pero las metáforas no son «sólo» metáforas; las metáforas son herramientas de pensamiento (1995, pág. 466)

⁶⁶I.e., que el dominio D pueda acomodarse en el ámbito de las ciencias naturales

Es curioso, sin embargo, que en todo el libro haya mostrado una innumerable cantidad de experimentos científicos, y sus conclusiones se basen en ellos, y le sirvan para dictaminar que la conciencia ha quedado *explicada*. Al final, sin embargo, parece que esa explicación es una “guerra de metáforas”. ¿Es que su explicación de la conciencia fue *metafórica*? ¿Debe entenderse a la ciencia natural como un conjunto de metáforas bien articuladas? ¿O algo más? ¿O algo diferente?⁶⁷

Nótese, adicionalmente, que el término “natural” es usado por tanto por filósofos dennettizadores como nagelizadores: por caso, David Chalmers, un filósofo típicamente “nagelizador”, asume que su posición es compatible con el naturalismo (Cf. Chalmers, 1999, pág. 19: “[L]a conciencia es un fenómeno natural sometido al dominio de las leyes naturales”. Págs. 19-20: “Todo lo que aquí digo es compatible con los resultados de la ciencia contemporánea; nuestra imagen del mundo natural se amplía, no se subvierte. Y esta ampliación permite una teoría naturalista de la conciencia [...]”).

Más que en las disparidades acerca de la posición naturalista, nos interesa remarcar el aspecto reductivo de la ciencia natural tal como la entienden tanto los dennettizadores como los nagelizadores. Para el dennettizador, recordemos, los fenómenos mentales pueden reducirse (en su caso límite, eliminarse) en favor de entidades naturales que pueden explicar todos los procesos mentales como eventos de un orden físico/funcional de bajo nivel. Para el nagelizador, esta estrategia es inaplicable debido a peculiaridades propias de la conciencia. Ambos, sin embargo, asumen que:

[A]lmost everyone concerned has been laboring under a mistaken assumption in thinking that reductive explanations are widespread within the natural sciences (2007, pág. 31)

Esta asunción les permite presuponer que es un hecho común, dado en todas las ciencias, que la estrategia reductiva funciona. Sin embargo es discutible que esto sea

⁶⁷ Colin McGinn en *Consciousness Evaded* (1995) tratará de establecer, a pesar de la afirmación de Dennett, qué debe entenderse como literal y qué puede ser metafórico. Según McGinn, es un hecho literal que: a) los animales y las personas poseen experiencia consciente; b) hay un sujeto unificado de las experiencias. De estos hechos, es necesario un entendimiento *literal*, no *metafórico*: “Is the generation of new metaphors what we need to make progress in understanding the mind? No, we need literal theories, such as the many now existence in psychology and associated disciplines” (1995, pág. 248)

así. El proyecto reductivista persigue el ideal de unidad reductiva: todas las ciencias podrían, en principio, ser reducibles a una ciencia fundamental, la física:

In the mid-twentieth century, the prevailing philosophical view was that the natural sciences themselves form a kind of natural hierarchy based on composition and complexity. The fundamental facts about the universe, as well as the fundamental laws, are those that are couched at the level of basic physics. (2007, pág. 23)

Este ideal se presumía extensible incluso a la psicología. Los nagelizadores dirán que tal ideal no puede cumplirse en el caso especial de la psicología, aunque sí puede concretarse en otros casos :

It is not absurd to suppose that psychological laws may eventually be explained in terms of the behavior of individual neurons in the brain; that the behavior of individual cells –including neurons- may eventually be explained in terms of their biochemical constitution; and that the behavior of molecules –including the macro-molecules that make up living cells- may eventually be explained in terms of atomic physics. If this is achieved, then psychological laws will have, in *principle*, been reduced to laws of atomic physics. (Putnam, “Unity of Science as a Working Hypotheses, 1958, citado en Horst, pág. 24)

Sin embargo, duante los años '80 del siglo XX surgen algunos de los más famosos argumentos acerca de la irreducibilidad de lo mental:

Levine (1983) claimed that there is an “explanatory gap” in the form of an irreducibility of several important mental phenomena to physical objects and processes. And this view was supported by several compelling thought-experiments, such as Jackson’s Mary, Nagel’s Bat, and Searle’s Chinese Room. Around the same time, Davidson (1970) argued on very different grounds that not only is the mental irreducible, but that there are not even any laws linking descriptions of the physical world to unique interpretations in terms of beliefs and desires (Horst, 2007, pág. 27)

El ámbito de lo mental se convierte en un caso especial que no puede recibir tratamiento reductivo. Pero de hecho no sólo no es el *único* ámbito, sino que las brechas explicativas y la imposibilidad de reducción son comunes en todos los terrenos de la ciencia natural. La reducción no se ha podido (ni se espera que se pueda en un futuro) operar en una enorme variedad de dominios. Existen serios problemas para efectuar reducciones interteóricas entre química y física, por ejemplo:

Often discussed cases of failed or incomplete intertheoretic reduction in the literature include:

- the reduction of thermodynamics to statistical mechanics (Primas

- 1991, 1998; Sklar 1999)
- the reduction of thermodynamics/statistical mechanics to quantum mechanics (Hellman 1999)
 - the reduction of chemistry to quantum mechanics (Cartwright 1997; Primas 1983)
 - the reduction of classical mechanics to quantum mechanics (such as the worry that quantum mechanics cannot recover classical chaos) (Belot and Earman 1997) (2007, pág.50)

También hay problemas de reducción en filosofía de la biología:

Biologists like Lewontin and Stephen Jay Gould (Lewontin and Gould 1979; Lewontin 1983) and philosophers like Philip Kitcher (1984) have argued that it is a misconception to view the relation between molecular biology and classical genetics or evolutionary theory as a reduction. First, while genes are a causal factor in determining phenotypic traits through development, they do not “determine” those traits. Every trait is a product of a combination of genetic inheritance and environmental influence through development. Phenotype is “plastic,” and sometimes quite dramatically so. Second, both the theory of evolution and developmental biology require us to view living organisms historically (over evolutionary and developmental time frames, respectively) and as systems that are open to their environments. (Horst, 2007, pág. 52)

Las brechas explicativas y los fracasos en la reducción son una nota común en la ciencia natural. Dado que la reducción (actual o futura) de los ámbitos científicos es lo que otorgaría unidad a la ciencia, la imposibilidad de reducir un ámbito a otro es síntoma de una *disunidad* en el conocimiento científico. Pero esta disunidad no parece ser un hecho epistémico que irá mostrando su *unidad oculta* con el avance de la investigación científica: la propia ciencia resulta ser disunida. El panorama es desolador si se adopta a la reducibilidad como un criterio de legitimación ontológica:

If reducibility is a criterion for ontological legitimacy, it is not only intentionality and phenomenology that are threatened, but phenomena of other sciences as well, such as organisms, species, metabolic processes, temperature, and molecules. (2007, pág. 80)

Más aun, la clausura causal del mundo físico no es, propiamente hablando, un hecho empírico sino una tesis filosófica. La clausura causal no puede derivarse ni de la verdad, ni de la totalidad de las leyes científicas. Esto implica que la “dirección reductiva”, de los fenómenos de alto nivel hacia los fenómenos de bajo nivel *no*

necesariamente ha de proveer todos los elementos nómicos necesarios para otorgar una explicación.

[O]ne must distinguish the following theses:

(C1): Laws L1, . . . ,Ln are true.

(C2): L1, . . . ,Ln are all the laws there are.

(C3): Events are causally closed under L1, . . . ,Ln.

Causal closure is established only in C3, and C3 is not entailed by either C1 or C2. It is an additional principle, one that must be added to a commitment to laws, and even to the philosophical premise that the domain of science is restricted to nomic phenomena. (2007, pág. 106)

La explicación reductiva solo puede funcionar si los hechos reducidos son explicados nomológicamente como hechos del nivel menor. Pero no puede garantizarse que todos los hechos causales involucrados sean *los únicos hechos posibles*, i.e., que haya clausura causal. Por lo tanto la estrategia reduccionista consistente en explicar los fenómenos mentales como parte de un universo explicativo que en sí mismo tiene cierre causal, falla incluso en la provisión del cierre causal que debería otorgar el propio ámbito al que se reduce el fenómeno mediante la explicación. En otras palabras, *aun cuando se opere la reducción, tal reducción no garantiza la provisión de la fuerza explicativa esperada.*

Estos diagnósticos de *disunidad científica, impotencia reductiva y ausencia de clausura nomológico-causal* nos proveen de un marco conceptual adecuado para entender por qué los casos de *parecer₁* examinados puntualmente en lo que hace a la visión del color, resultan tan intratables: la *disunidad* del color (tanto en su aspecto fenoménico como en su aspecto funcional) es probablemente un síntoma de la *disunidad* de la ciencia; la brecha explicativa entre fenómeno y acaecimiento es una de las muchas brechas de la ciencia. Cuando Locke apela a la arbitrariedad divina como expediente no nomológico para explicar la brecha entre las ideas de cualidades secundarias y los hechos del mundo, está aceptando esta disunidad. Horst propone una visión de la ciencia a la que denomina *pluralismo cognitivo*, una posición que no persigue la búsqueda de reducción interteórica ni espera la unificación de todos los campos científicos:

The position I wish to outline here I call Cognitive Pluralism. It is “cognitivist” in that it traces features of our understanding of the world to features of our cognitive

architecture, that is, to empirical facts about how minds like ours model features of the world. It is also, in some ways, a Pragmatist position, in that the nature of our explanatory interests and of our interactions with the world also plays a role in determining the form of models in science and outside it. It is a pluralist position in that it holds that we relate to the world through an irreducible plurality of special-purpose models that are not reducible to a single common denominator or unifiable into a single axiomatic system. (...) we do not have anything so global as a worldview at all. Instead, we triangulate the world by deploying various models, each of which is good enough for particular things. (...) However, the partial and idealized character of these models, *qua* cognitive models, poses barriers to their integration into a single supermodel. (2007, págs. 127-8)

La ciencia utiliza una variedad de modelos adecuados a propósitos diferentes, y resulta imposible articular todos los modelos en un solo supermodelo unificado. Los modelos son idealizaciones, y como tales, sólo pueden conceptualizar determinados aspectos del mundo: *explicar es dejar de lado algunos aspectos del hecho a explicar*. Estos modelos son parcialmente determinados por nuestra arquitectura cognitiva, y cada uno de ellos describe determinados aspectos del mundo. Por último, cada uno de los aspectos anteriores determina que los modelos no puedan unirse en última instancia en un supermodelo. Horst resume esto en seis puntos:

1. Scientific laws and theories are models of particular aspects of the world.
2. Such models are products of cognitive processes of modeling, and hence their form is determined in part by human cognitive architecture.
3. Scientific models are idealized, and indeed can involve several forms of idealization.
4. Each model must employ some particular representational system in describing its subject matter.
5. Both the idealization of models and the choice of representational systems can present barriers to the integration of different models and may necessitate the partial rather than comprehensive character of individual models.
6. Empirical facts about human cognitive architecture will constrain the types of models we can conceive, understand, and employ, and there can be disunities within the sciences that are artifacts of our cognitive architecture. (2007, págs. 128-9)

Podría pensarse que distintos dominios de hechos del mundo no pueden confluír en un único esquema representacional debido a la disparidad de modelos. Pero incluso *un único* tipo de hechos requiere de modelos explicativos diferentes, que se interfieren mutuamente, y cuya “modelización” en conjunto es impracticable (quizás esta era una razón por la cual resultaba tan problemático encontrar una conceptualización general

de la visión, tanto en su aspecto fenoménico como en su aspecto funcional: la visión parece necesitada de más de un modelo explicativo). Por ejemplo, la ley de gravitación universal como modelo para explicar la caída de los cuerpos:

[M]ay in some sense “have universal scope” in that it is always “in play”; but what it does not do is tell us how things always actually behave in nature. Objects do not actually behave as the gravitation law would predict when there are other causal factors, such as aerodynamics or electromagnetism, at work as well. A paper airplane does not fall like an identical piece of paper of equal mass crumpled into a tight ball. A metallic object does not fall if it is in the field of a suitably strong magnet. Even laws like the gravitation law, if taken as universal claims about how objects actually behave, have exceptions. Indeed, because multiple forces are always “in play,” it is likely that even many of the most basic laws have nothing but exceptions, if interpreted as universal claims about actual behavior. (2007, pág. 76)⁶⁸

La utilización de un único modelo (por ejemplo, el de la gravitación universal) para explicar ciertos hechos puede ser suficiente para determinados propósitos en los que no se requiere de una explicación (o una predicción) fina o conceptualmente adecuada (CAE). Pero incluso para hechos cotidianos, la explicación conceptualmente adecuada de ciertos sucesos (por ejemplo, la trayectoria de un avión de papel arrojado al aire) sólo puede realizarse si se utiliza más de un modelo explicativo (leyes de aerodinámica y electromagnetismo, además de las leyes gravitatorias). “Indeed, even in the cases in which a gravitational model alone is good enough for practical purposes, the results of plugging initial conditions into that model never predict real-world behavior with complete accuracy, as other forces are always in play.” (2007, pág. 104). El comportamiento en el mundo real debe ser producto de una completa suma de fuerzas, “iterated over real time, over the initial conditions. But sometimes there is not a technique for summing forces. And the disparate physical models we employ to explain different parts of a phenomenon are sometimes idealized in incompatible ways; for example, as Mark Wilson (2006) points out, crucial terms like ‘force’ may

⁶⁸ Haremos una salvedad a esta cita: no compartimos el criterio de que la mayoría de las leyes no tienen sino excepciones. En realidad, lo correcto sería decir que las leyes como la de la gravedad actúan siempre, solo que en determinadas escalas son interferidas por otras leyes que también están actuando. Que un avión vuele no es una excepción a la ley de la gravedad. En rigor, una auténtica excepción sería un milagro. Lo que deseamos enfatizar con esta cita es que resulta sumamente difícil (por no decir impracticable) predecir el comportamiento de un cuerpo en movimiento dada la cantidad de leyes que interactúan y la cantidad de modelos no unificables que se requieren para lograr una predicción precisa.

have divergent and incompatible meanings within the context of different models.” (2007, pág. 104)

Los modelos aplicados por lo general subdeterminan el comportamiento de los hechos en el mundo real. “And on the other hand, the idea that resultant behavior is determined by ‘a summation of all the forces that are in play,’ which looks so attractive in the abstract, often proves unworkable or even unintelligible when one turns to the details of the models employed by the scientist or the engineer” (2007, pág. 104). Horst afirma que esta imposibilidad no es un simple síntoma de inmadurez científica, sino que se trata de una sistemática y generalizada discordancia en *todos* los ámbitos científicos.

Horst analiza tres diferente tipos de idealizaciones que proporcionan los modelos: la idealización “ignorante” (“bracketing”), la “distorsionadora” y la “aproximativa”.

- a) La idealización “**ignorante**” (bracketing) supone que los modelos sólo pueden modelar una *parte*, aunque no el *todo*. Por eso, el modelo es semejante a una captura momentánea, de un determinado aspecto y en un determinado momento, pero no del hecho *en tiempo real*. Los aspectos no modelados simplemente se ignoran; se dejan de lado en favor de otros aspectos que se consideran más relevantes o mejor tratables. Por supuesto, en muchos casos (como en el caso de un modelo gravitatorio operante en un avión de papel arrojado al aire en el cual se *ignorasen* modelos de dinámica y electromagnetismo) las variables que se ignoran *están de hecho operando de modo relevante*. En todos los casos, el sistema real –a diferencia del modelo- no es independiente de los hechos ignorados. (Horst, 2007, págs. 131, 134)
- b) La idealización “**distorsionadora**” ocurre cuando el modelo ofrece una representación distorsionada para ajustarla a ciertos parámetros ideales. Por ejemplo, en economía, las personas son tratadas como *agentes racionales* ideales, y en un modelo de la colisión de bolas de billar se presupone que las bolas no se deforman con los golpes: se las trata como *cuerpos perfectamente elásticos*. (Horst, 2007, pág. 135) “Rather, the point (or one point, at any rate)

is that we have no single general model of collisions, but cover the various cases through a variety of ‘patches.’” (2007, pág. 137).

- c) La idealización “**aproximativa**”: el modelo utiliza constantes aproximadas, no exactas. Tal es el caso del valor adoptado por π para un determinado modelo: es una decisión puramente pragmática dónde deben truncarse los decimales: “3,14” o “3,141592”. (2007, págs. 143 - 4)

Estos tipos de idealizaciones presentes en los modelos no permiten una reflexión unitaria y precisa del fenómeno del mundo real que están modelando. Si para la explicación de un hecho se adoptan *varios* modelos que idealizan cada uno de modo diferente, resulta imposible compatibilizar los tipos de idealización para lograr una única explicación convergente.

El pluralismo cognitivo adoptado por Horst propone que cada ámbito del mundo debe ser modelado de acuerdo a nuestro interés *práctico*, el cual difiere según el tipo de hecho que se pretende explicar. De aquí que su propuesta sea fundamentalmente pragmática. El interés *teórico* es una subcategoría de este interés práctico, y cuáles son nuestros intereses es *un hecho acerca de nuestra propia mente*: nuestra arquitectura mental nos provee de ciertas demandas cognitivas y no otras. Por eso, para determinar por qué hay disunidad en la ciencia, es necesario determinar cuáles son las limitaciones y peculiaridades de la cognición humana. Es posible que el trabajo cognitivo humano funcione mejor si se realiza moviéndose *a través* de modelos diferentes de acuerdo con la demanda pragmática, antes que buscando un modelo unificado, omnicomprendido e independiente de los intereses prácticos (2007, pág. 140). Nuestra arquitectura cerebral está diseñada mediante una *división del trabajo cognitivo*. Horst afirmará que:

1. Cognitive division of labor is a deep and ecumenical design principle in human and animal brains.
2. At least one form this takes is in the creation of functional processes in the brain that model aspects of the organism, its environment, and relations between the two.
3. This division of labor can be realized in a number of ways, including but not limited to the localization of cognitive function in neural areas and layers.
4. We may see the extension of cognitive division of labor, from the proliferation of special-purpose areas to neurally distributed (though perhaps not fully global)

special-purpose acquired abilities (...), as an instance of a single design principle that is implemented in different mechanisms and on different timescales: through gene selection resulting in special-purpose brain areas; through redeployment of existing areas in new functional configurations; and through learning in neural networks optimized to form pragmatically driven, interest- and organism-specific, idealized models.

5. Scientific models are special cases of these latter capacities. (Horst, 2007, pág. 154)

Dada la arquitectura de trabajo especializado de los diferentes mecanismos del cerebro, podemos encontrar una clara división modular de tareas cognitivas. En lo que respecta a la visión, existe un detallado y complejo proceso que, a pesar de ser un fenómeno de apariencia unitaria, está formada por una multitud de elementos modulares independientes:

[W]e intuitively think of “seeing” as a single mental process— Hume suggested the metaphor of understanding it as a drama played out on a single imagistic stage, for example. Yet information from the retina is quickly split into three separate pathways for boundary/shape, color, and motion, which are registered in different layers of cells in the LGN and separate regions of the visual cortex (Felleman and van Essen 1991). All three streams of information pass through the LGN, V1, and V2, albeit through different layers of cells. V1 and V2 perform a great deal of the work in extracting information about form.

Information about motion is passed on from V1 and V2 to V3 and V5, and information about color to V4. From the visual cortex, information is further split into a “what” pathway (located ventrally in the temporal cortex) and a “where” pathway (located dorsally in the parietal area). The “what” pathway is involved in recognizing types of objects and seems even to include such exotica as an area (the fusiform gyrus) specifically devoted to the recognition of faces. The “where” pathway is involved in both locating objects in space and in interacting with them kinesthetically. Such cerebral division of labor seems to be the rule rather than the exception in the human brain. (2007, pág. 155)

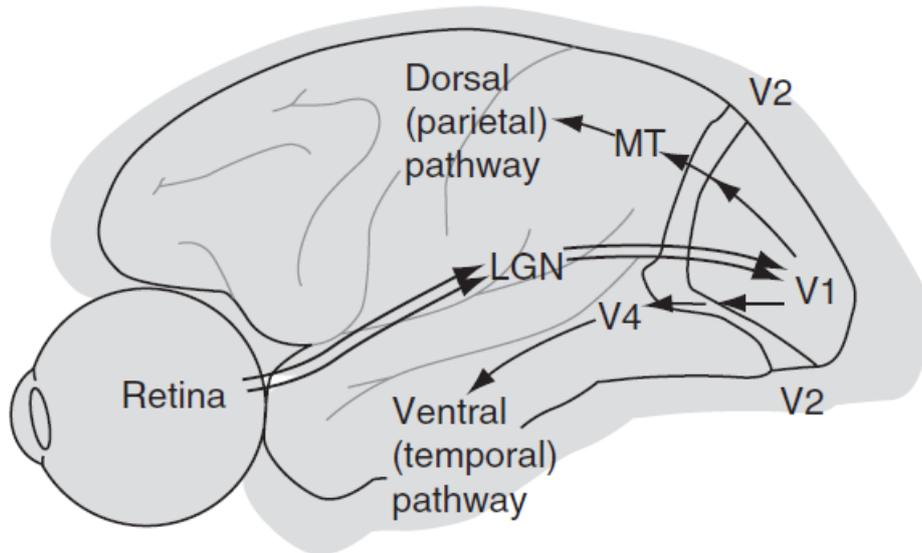


Figura 13. (Extraída de Horst, 2007, pág. 156)

Así como existe una división de trabajo en lo que respecta a la estructura funcional que provoca los fenómenos de la visión (tal como afirma el punto 3), también existe esta división en nuestras capacidades cognitivas en general y, en consecuencia, en los diferentes tipos de modelos científicos con los que intentamos entender el mundo (punto 5). Los modelos científicos no pueden escapar de esta arquitectura modular del procesamiento mental. Sin embargo, “modular” es un término que envuelve, al menos, dos contextos semánticos: a) modularidad puramente cognitiva, sin específica localización neural. En este caso, la modularidad se entiende como un proceso encapsulado, en el cual otros procesos no pueden inherir. Cada módulo está “cognitivamente encapsulado” y no es afectado por la acción de otros módulos; por ejemplo, en el caso de las ilusiones ópticas analizadas más arriba (v. § 2.3), el anillo rosado y el círculo verde no desaparecen aunque el individuo que las percibe conoce que se trata de una ilusión (Fodor, 1983, 68, citado en Horst, 153). La modularidad funciona como un encapsulamiento cognitivo (la visión del anillo rosa no se ve “afectada” por la información recibida desde otros módulos cognitivos, como el de la creencia), aunque no se postula una localización neural de tales módulos; b) al contrario de la acepción anterior, el otro sentido de “modular” se refiere a módulos cerebrales específicos para cada tarea cognitiva. Existe una evidencia fuerte de que

muchas tareas cognitivas están anatómicamente localizadas en el tejido neural, y por lo tanto parece plausible asumir la acepción b). Además del ejemplo de la visión puesto más arriba, existen innumerables ejemplos en los que las funciones psicológicas tienen una correlación fuerte con áreas cerebrales: “it was observed that patients with highly localized brain injuries characteristically experience loss of narrowly defined psychological capacities as well, and that there are many robust correlations between brain areas and psychological functions” (2007, pág. 154) La división de los procesos cognitivos en una multitud de áreas cerebrales es un diseño básico de la arquitectura cognitiva de los mamíferos.

Sin embargo, no todas las funciones pueden localizarse anatómicamente. Por un lado, existen habilidades cognitivas que no pueden identificarse con un área cerebral. Por ejemplo, la habilidad para jugar al ajedrez o la “localización cerebral” de una palabra. Por otro lado, las distintas áreas del cerebro poseen una enorme plasticidad para asumir funciones diferentes en caso de daño de alguna de las áreas. Pero, aunque no puede encontrarse una localización anatómica de ciertas funciones, es posible que exista un “masivo redespliegue” (*massive redeployment*) de áreas cerebrales existentes para lograr un nuevo tipo de funcionalidad. (2007, pág. 165). Puede que no haya una relación uno-a-uno entre una función cognitiva y un área cerebral, precisamente porque la relación es jerárquicamente compleja, como se ilustra en la figura siguiente:

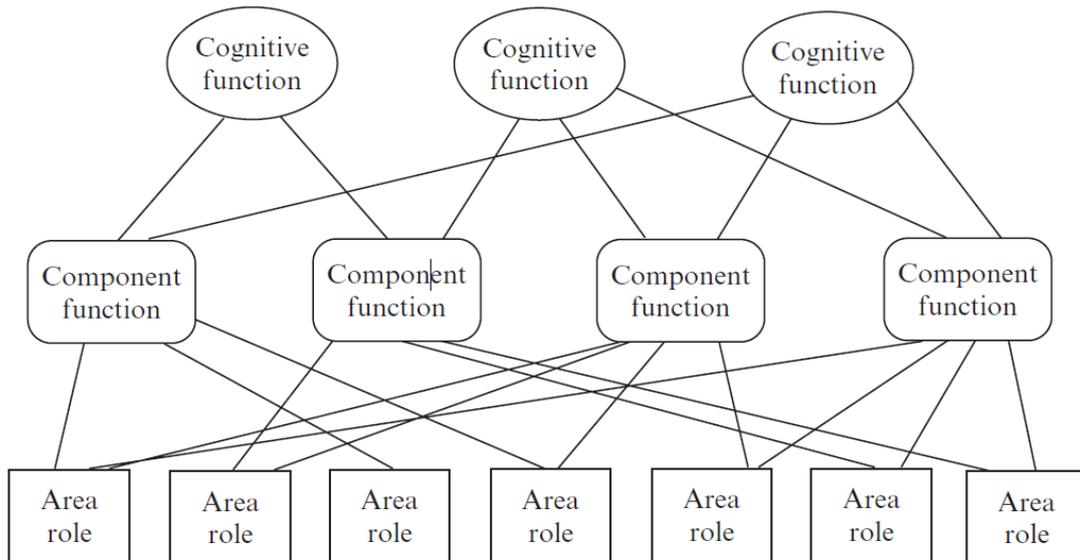


Figura 13. (Extraída de Horst, 2007, pág. 165)

Las actividades cognitivas de nivel superior no tienen un *locus* cerebral asignado; son más bien una consecuencia de la actividad conjunta de varias áreas no necesariamente contiguas. Su actividad es más parecida a la de un pelotón o un comité que a la de un individuo (2007, pág. 166).

Si se acepta que hay una división de tareas cognitivas, existe un problema adicional que no posee una respuesta definitiva: cómo los diversos componentes cognitivos se “atan” para producir una experiencia unificada (*the binding problem*). Si existen áreas separadas para procesar la representación del movimiento, el color y la forma, en la visión de un triángulo rojo y un círculo azul “we need a story about «how the brain knows that ‘red’ goes with ‘triangle’ and ‘blue’ goes with ‘circle.’» Why don’t we instead link ‘red’ with ‘circle,’ or with both ‘triangle’ and ‘circle,’ or indeed with no shape at all?” (2007, pág. 167) Daniel Dennett (1995) afirma que no existe un módulo superior que una las representaciones (un “teatro cartesiano” en el que se “proyectan” las representaciones para que “el sujeto” las perciba de manera unificada), porque no existe esa experiencia unificada. Horst propone, al contrario de Dennett, que si bien no existe una “representación comprensiva” en un módulo superior, sí existe una unificación de la experiencia en el sentido de que los espacios intermodales no *parecen*₁ separados en una experiencia sensorial; por ejemplo, en la experiencia de

una botella de color verde. Aunque no se postule un módulo de representación ulterior, sí puede postularse la unidad de la experiencia en tanto a una variedad de representaciones dependientes de módulos separados se las experimenta como *la experiencia de un único objeto*:

Denying that there is a unified representational system for form and color in the brain does not require us to deny that there are other forms of unity present: in the case at hand, the kind of unity involved in constituting something as a single object with multiple properties, though one might add to this other types of unity, such as crossmodal identifications involving sound, sight, and touch, constituting successive experiences as experiences of the selfsame object (2007, pág. 168)

Es importante recalcar que la unidad de un objeto no es una unidad representativa, en el sentido de que dicha unidad no viene dada por una representación ulterior (por la activación de un módulo representacional omnicomprendivo), sino que la sucesión o la simultaneidad de las diferentes representaciones constituyen la unidad de la experiencia, aunque esa unidad no está dada *qua representación de unidad*. Este modo de proceder de la mente es aplicable a los modelos científicos, en tanto que la unidad científica está dada por la unidad de la naturaleza (una unidad que recuerda al espíritu de la *idea de mundo* kantiana⁶⁹), los modelos científicos no pueden representarse a la naturaleza en un único esquema representacional. “While the cognitive pluralist, by definition, resists the assumption that *representational unity is the norm*, she need not resist the idea that there are important, if partial, connections between models of regional domains, nor deny that we can deploy them in tandem to get a better understanding than we can from any of them taken individually” (2007, pág. 172). Aun cuando no pueda haber un esquema representacional unificado, diferentes modelos pueden “trabajar en conjunto” para lograr una mejor representación de un hecho. Sin embargo, no puede esperarse que esta representación sea unificada en una representación ulterior, última y omnicomprendiva.

El caso del problema de la relación entre fenómeno mental y acaecimiento es un caso más entre otros, y no posee un status especial, pues la explicación de esa relación utiliza modelos que son producto de nuestra especial arquitectura cerebral.

⁶⁹ “[W]e might recall Kant’s cautions about the dialectical illusions of the Ideas of Reason” (Horst, 2007, pág. 180)

Podemos resumir algunas consecuencias del análisis de Horst para la explicación de los fenómenos mentales:

1. No puede haber un único modelo representacional del mundo.
2. La mente posee una arquitectura tal que sólo puede poseer múltiples modelos para entender al mundo, cuya unidad es un postulado ideal.
3. No es posible elaborar un único modelo para entender el diseño y el funcionamiento de la propia mente.
4. Cualquier modelo explicativo unitario que se proponga explicar *sin resto* la naturaleza de la conciencia, tiene la pretensión de capturar los hechos mentales en un único modelo; i.e.: persigue el ideal de unidad de los diversos modelos.
5. Explicar la naturaleza de la mente es utilizar la mente (que posee una determinada arquitectura cognitiva) para realizar un modelo representacional de la mente.
6. Pero en el acto de entender a la mente bajo un único modelo se soslaya el hecho *acerca de la propia mente* de que la propia arquitectura cognitiva está disunificada; i. e., un único modelo puramente teórico se muestra incapaz de capturar la esencia de lo mental. (Los modelos teóricos se elaboran para propósitos prácticos)

De los puntos anteriores, nosotros podríamos concluir que:

7. Es posible que una estructura explicativa reduccionista deba combinarse con una estructura explicativa fenomenista (o, en términos que hemos usado anteriormente, que una explicación dennettizada deba combinarse con una explicación nagelizada), aun cuando los modelos generados por estos tipos de explicación sean incompatibles, pues:
8. Es un hecho acerca de la arquitectura cognitiva humana que no puede esperarse una compatibilidad ulterior, última, de los modelos empleados. Por lo tanto, es plausible construir un modelo de explicación de los hechos de conciencia que sea fenomenista (“nagelizador”) y otro que sea reduccionista (“dennettizador”), de acuerdo a los propósitos de la explicación, aunque no es esperable que uno de ambos modelos deba erigirse como *el* modelo de la conciencia.
9. En la explicación unificada de la conciencia no puede desconocerse la arquitectura cognitiva humana, ni la fenomenología (en tanto esta fenomenología está presente en el parecer₁) de lo mental. Pero una explicación no tiene por qué –ni es deseable, ni es posible– elaborar un modelo que satisfaga *a la vez* ambos aspectos.

10. La explicación dennettizadora es legítima, aunque no puede esperar el logro de una reducción total y unificadora de todos los fenómenos mentales, sino diversos tipos de idealizaciones parciales. A su vez, la legitimidad de la explicación dennettizadora no tiene por qué dejar fuera de juego a la explicación nagelizadora. Que exista (o que se busque) un conjunto de modelos explicativos funcionales no unificables para los fenómenos de la conciencia no implica que la demanda nagelizadora de una explicación del fenómeno *qua* fenómeno deba ser irrelevante.

La mente despliega una multitud de modelos para entender al mundo. En tanto busca entenderse a sí misma como parte del mundo, pretende subsumirse a sí misma bajo uno de los modelos explicativos pertinentes para ciertos aspectos del mundo. Pero, como los metafísicos de la nación imaginaria de Tlön, quienes persiguen este objetivo “[s]aben que un sistema no es otra cosa que la subordinación de todos los aspectos del universo a uno cualquiera de ellos” (Borges, 1999, pág. 18) Un modelo de la mente es el intento de subsumir a la mente (que despliega multitud de modelos) bajo uno de esos modelos. Aun cuando la subordinación no sea a *uno cualquiera* (la elección del modelo no es arbitraria, sino que obedece a propósitos pragmáticos), el ideal de unidad que está en la base del modelo reduccionista pretende excluir a otros modelos en tanto explicativamente relevantes, y se presenta a sí mismo como *el* modelo científico. “Scientific models are a special case of mental models generally” (2007, pág. 172): La mente genera modelos; algunos de ellos son modelos científicos acerca de hechos del mundo; algunos de ellos son modelos científicos *de la propia mente*. El instrumento empleado para capturar la esencia de lo mental es un modelo mental, como un cuadro que se pinta a sí mismo o un ojo que, en el acto de ver, pretende verse a sí mismo. Existe, sin embargo, un peligro del que debemos alejarnos: decir que “la mente genera modelos” no significa, en rigor, que los modelos sean puramente mentales o incluso individuales. La ciencia es una labor colectiva que transcurre en el tercer mundo de lo conceptual, no en el segundo mundo de lo mental⁷⁰. Pero los modelos de ese tercer mundo, aun como modelos científicos colectivos, tienen ciertas

⁷⁰ Nos mantendremos neutrales con respecto al problema ontológico del “tercer mundo”. ¿Es un reino de los sentidos que se capta mediante una actividad del pensamiento, como pensaba Frege, o mas bien es un conjunto de construcciones simbólicas creadas por el hombre? Si tomamos la posición platónico - fregeana, quizás nuestra mente solo pueda captar modelos parciales y no unificables, limitación para la cual tal vez debamos, de todos modos, encontrar una explicación en la arquitectura cognitiva humana.

características que parecen inherentes a la arquitectura cognitiva de los sujetos para quienes esos modelos son modelos:

[C]haracteristic features of models: they are partial, idealized, and pragmatically driven, and they employ proprietary representational systems. These factors alone—the features of models as models—are capable of creating barriers to integration of models, including irreducibility and inconsistency. Thus, the hypotheses that human thought is model-based, and that scientific thought is a special case of model-based thinking, provide candidate explanations for the disunities of both ordinary thought and the sciences. They also entail at least the possibility that such disunities are an unavoidable artifact of our cognitive architecture. (Horst, 2007, pág. 175)

Hay quizás algo en la arquitectura mental humana que nos lleva a representaciones intersubjetivas del mundo que son ulteriormente imposibles de unificar o de reducir. En ese sentido, el problema de la unidad y de la reducibilidad de lo mental no es un problema especial, en tanto las brechas explicativas se encuentran en todos los ámbitos de la ciencia.

Conclusiones: ¿Dónde está la conciencia? Un diagnóstico especulativo

Quizás, como concluimos en el apartado anterior, haya algo intrínseco en la arquitectura humana que no permita una unificación ni una reducción total de los modelos utilizados, tanto en lo que respecta a la conciencia como en la mayoría de los ámbitos de la ciencia. Sin embargo, muchas veces la esperanza reduccionista se vale de ciertas analogías que buscan correlacionar lo mental con un hecho físico o perceptivo. Sostendremos que estas analogías están mal construidas; son fallidas en su intento de mostrar que la mente es análoga a otro proceso porque en ese otro proceso deben presuponer a la mente.

Tomemos en primer lugar la siguiente analogía wittgensteiniana presentada en las proposiciones 5.633 y 5.6331 del *Tractatus* de Wittgenstein (1999). En rigor Wittgenstein no buscaba una estrategia reduccionista, pero nos sirve como ejemplo inicial para visualizar la falla en la analogía:

5.633: “¿Dónde descubrir en el mundo un sujeto metafísico? Dices que ocurre aquí enteramente como con el ojo y el campo visual. Pero el ojo *no* lo ves realmente. Y nada en *el campo visual* permite inferir que es visto por un ojo”.

5.6331: “El campo visual no tiene, en efecto, y por así decirlo, una forma como ésta:”

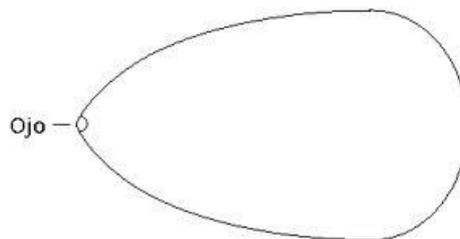


Figura 14. (Extraída de Wittgenstein, 1999, pág. 145)

El ojo no se puede ver a sí mismo en el acto de ver, del mismo modo que el sujeto metafísico no puede encontrarse en el mundo. En tanto sujeto metafísico, es conciencia del mundo, pero no es parte del mundo. Sin embargo, el ojo podría verse a

sí mismo en un espejo. Podría, también, observar una cámara en la que se viera a sí mismo viéndose a través de la cámara. Podría, si la cámara estuviera conectada con algún sistema de escaneo cerebral, ver una representación de las áreas corticales que se activan en el acto de ver la cámara que le muestra las representaciones de las áreas corticales que se activan en el escaneo cerebral. Podría ver todo esto pero, como dice Wittgenstein, no tiene por qué inferirse que eso que se está viendo está a su vez *produciendo* la visión. Podemos objetar que esta historia tampoco es totalmente cierta: el ojo, en sí mismo, no *produce* la visión⁷¹. No es el ojo el que ve, sino el sujeto consciente quien ve a través del ojo. Si no hay un sujeto consciente (en términos de Frege, un *portador* de las representaciones) no hay, en ningún sentido relevante, una representación. La comparación del ojo que no puede verse a sí mismo no es una buena analogía de la mente que no puede encontrarse en el mundo, pues el análogo utilizado (el ojo que no se ve) ya presupone a la propia conciencia (el analogado). El análogo se supone que es mejor conocido que el analogado. Se supone que el analogado es diferente al análogo (si no lo fuera, no tendría sentido establecer la analogía, pues habría identidad entre el análogo y el analogado), pero comparte una misma propiedad con este último⁷². La analogía del ojo tiene la siguiente estructura:

- 1) *La mente no puede encontrarse en el mundo, (analogado) del mismo modo que:*
- 2) *El ojo no puede encontrarse en el campo visual. (Análogo)*

Pero: el campo visual presupone a la mente. Por lo tanto, la mente aparece tanto en 1 como en 2; i.e., la mente (su imposibilidad de ser encontrada en el mundo) aparece tanto en el analogado como en el análogo (el campo visual, que es generado por una mente). La analogía funcionaría si se nos dijera que la mente tiene tal o cual imposibilidad, *del mismo modo que otro hecho diferente* tiene una imposibilidad similar. Pero si ese “otro hecho diferente” esconde el primer hecho del cual

⁷¹ Afirmar lo contrario sería caer en una modalidad de la falacia mereológica. Esta falacia se comete cuando se toma una parte por el todo. En el caso señalado, se supone que el ojo es necesario y suficiente para caracterizar la etiología de la visión. Pero el ojo solo es una parte de esta etiología y, en rigor, no puede decirse que es el ojo el que ve, sino la persona. Para un examen de esta falacia con relación a los atributos psicológicos, véase Bennett *et al.*, 2008

⁷² La distinción entre ‘análogo’ y ‘analogado’ puede encontrarse en Pereda (1994), págs. 270-7

buscábamos una analogía, entonces la analogía falla. Lo que 2) quería mostrar ya estaba mostrado en 1).

Es curioso que en la historia de las explicaciones acerca de lo mental encontramos una cierta cantidad de analogías fallidas de este tipo, en las que el analogado (la conciencia o alguna característica de esta) debe presuponerse en algún sentido importante para interpretar el análogo. Por ejemplo, Daniel Dennett afirma que el yo es *fictional*; se trata de un centro de gravedad narrativa (1995, cap. 8): al igual que con Wittgenstein, la comparación no funciona, porque cualquier mundo ficcional sólo es ficcional para un yo. ¿Es el “yo ficcional” de un personaje de ficción como Sherlock Holmes (el análogo) *del mismo tipo* que mi propio yo (el analogado)? De algún modo, el análogo parece presuponer al analogado, pues es el yo el que crea ficciones (y para el cual las ficciones cuentan *como* ficciones: Sherlock Holmes sólo puede ser un yo si existe, en algún sentido, un “yo real” que le dé sustento como “yo ficticio”), de modo que no podemos comparar a la ficción con el yo (lo ficcional sólo tiene sentido si existe un sujeto para el cual lo ficcional cuenta como tal). Collin McGinn, en “Consciousness evaded”, (1995, pág 246) duda del valor de esta caracterización del yo como ficción: “Taken literally, this has to be false, of course, on pain of not distinguishing real people from fictional ones. But one assumes that the talk of fiction here is itself metaphorical or non-literal. Perhaps it is only fictionally true that people are fictions, so that literally they are real!” Pero si entendiéramos al yo como algo ficcional *de modo literal*, ¿cómo algo ficcional puede generar ficciones? ¿Es la ficción inicial del mismo tipo que las ficciones generadas por esa ficción? Una “ficción agente”, ¿no oculta después de todo a un agente separado de esa ficción? En algún momento, alguno debe no ser ficticio. ¿Puede *el propio Sherlock Holmes* crear un mundo ficticio? Como en el caso de Wittgenstein y de Frege, es necesario asumir una conciencia (un portador) que tome las riendas de estas ficciones.

Las explicaciones reduccionistas acerca de la naturaleza de lo mental parecen contener algún tipo de analogía fallida como las analizadas anteriormente. Que el analogado aparezca en el análogo indica que hay algo redundante: lo que queríamos mostrar con la analogía, ya estaba presupuesto en el análogo. Por esta razón, podemos llamar a estas analogías “Analogías QED (*quod erat demonstrandum*)”.

Los dos siguientes argumentos analógicos son del tipo *analogías QED*:

1. *“Con la conciencia y el cerebro ocurrirá lo mismo que con las dos mesas de Eddington⁷³. Mientras la mesa cotidiana es sólida, tiene extensión y es posible sostener un papel con ella, la otra –la mesa científica- es un enjambre de cargas eléctricas que se precipitan a gran velocidad. La mente es a la mesa cotidiana lo que el cerebro es a la mesa científica”*. Pero la mesa cotidiana (el análogo) es una construcción conceptual que involucra la percepción de los sentidos, los cuales presuponen a la mente (el analogado). Cuando se opone “solidez de la mesa” a “enjambres de cargas eléctricas”, se quiere decir: “la experiencia de solidez para un sujeto” versus “la explicación científica”. En la primera parte (el análogo) se habla de una experiencia que involucra la conciencia. Si esto es así, la analogía “la conciencia es al cerebro (analogado) como la solidez es a la mesa (análogo)”, debiera explicitarse del siguiente modo: “la conciencia es al cerebro (analogado) lo que la *experiencia consciente* de una mesa es al movimiento de electrones (análogo)”. Como puede verse, la conciencia aparece *dos veces*, tanto en el analogado como en el análogo. Y para que la analogía sea provechosa, necesitamos que la conciencia sólo aparezca en el analogado, pues es acerca de la propia conciencia de quien necesitamos saber cuál es su relación con el cerebro.

2. *“Así como la aparente inmovilidad de la Tierra puede explicarse a partir de modelos gravitatorios y celestes (en los cuales la inmovilidad desaparece) [análogo], la mente podrá explicarse a partir de algún modelo en el cual la propia fenomenología desaparezca [analogado]”*. Pero la “aparente inmovilidad de la Tierra” del análogo es un fenómeno consciente. Una vez más, la consciencia aparece *dos veces*, tanto en el análogo como en el analogado.⁷⁴

⁷³ En Eddington, *La naturaleza del mundo físico*.

⁷⁴ Aquí la analogía falla, además, porque es el movimiento lo aparente, mientras que las entidades involucradas (los astros) no lo son. La apariencia se clarifica cuando las entidades son reubicadas o reorganizadas. El problema con la “mente” es que las entidades, “los fenómenos mentales”, no se reconfiguran, sino que se reducen o se eliminan.

Estas analogías se suelen utilizar para mostrarnos la plausibilidad de una reducción entre ámbitos en apariencia disímiles: si algunas propiedades de la conciencia son, en algún sentido, semejantes a algunas propiedades de otro hecho no mental cuya explicación reductiva ha sido exitosa, entonces es esperable que en el ámbito mental la reducción tenga el mismo éxito. Pero si el análogo utilizado contiene un elemento fenoménico, entonces la reducción no podrá operarse.

En el transcurso de este trabajo nos hemos encontrado algunas aplicaciones de analogías QED. En lo que respecta al fenómeno de la visión, Jeff Foss (cf. § 2.5) y Adam Pautz (cf. § 2.6) proponían que la teoría de pares oponentes constituía el modelo funcional de las relaciones entre colores. La analogía se podía establecer del siguiente modo: así como existen relaciones de similitud y diferencia entre la experiencia fenoménica del color (analogado), existen correlaciones similares en el sistema de pares oponentes (análogo funcional). Pero tanto Mollon como Byrne & Tye (§ 2.6) argumentaban que el sistema de pares oponentes es un esquema fenoménico. De modo que el analogado aparece en el análogo.

Cuando analizamos la noción de *parecer* en Dennett (§ 2.3), encontramos un argumento que contenía la siguiente analogía: así como nos parece que hay un anillo rosa (análogo), nos parece que hay fenomenología (analogado). La intención de Dennett era negar la fenomenología pero, como señalaba Colin Mc Ginn (§ 2.3) parece haber una contradicción en afirmar un “parecer” en la misma proposición en que está negando el carácter fenoménico. De modo que la fenomenología aparece, usada en el mismo sentido (como el parecer₁), tanto en el análogo como en el analogado.⁷⁵

Ocurre algo similar con la explicación de Jeff Foss y Adam Pautz acerca del fenómeno de la visión: así como los fenómenos de la visión en colores tienen una estructura

⁷⁵ La analogía funcionalista (que forma parte de la tesis débil de la inteligencia artificial; cfr. Searle, *Minds, Brains and Programs*, [1980]) según la cual “La mente es al cerebro [analogado] lo que el *software* es al *hardware* de una computadora [análogo]” no comete, estrictamente, esta petición de principio, pues en el análogo no aparece la mente, sino el *software*. En realidad esta analogía difiere de las expuestas más arriba, pues no necesariamente apunta a la posible reducción entre mente – cerebro por un lado y *software* – *hardware* por el otro: dice que en importantes puntos son similares o parecidos, pero no está haciendo afirmaciones con respecto a la reductibilidad de uno a otro (es problemático, también, entender qué significa que el *software* se reduzca al *hardware*: esto es, la reductibilidad no resulta evidente ni siquiera en el análogo). Un programa es un sistema de símbolos, y no es fácil entender de qué modo ese sistema puede “reducirse” al hardware que le sirve de soporte. Para algunos lineamientos clásicos de esta discusión véase Dennett, (1987) y Putnam, (1960)

determinada en la cual ciertas combinaciones son posibles y otras son imposibles –por ejemplo, verdirrojo- (análogo), esa estructura es correlativa con la teoría de los pares oponentes en la cual solo pueden darse ciertas combinaciones de acuerdo al movimiento de un *switch* en el perímetro de las combinaciones posibles (analogado). Pero, como vimos más arriba, la teoría de los pares oponentes es ella misma una teoría fenoménica.

Las peticiones de principio presentes en estas analogías parecen proceder de un ancestro común: el marco cartesiano de la relación entre las ideas y el mundo. En todos estos casos, se desea dar una explicación de los estados fenoménicos, internos, a partir de elementos no fenoménicos, externos. Pero el aspecto externo no puede formar parte del aspecto interno o, en otras palabras, los fenómenos (ideas / pareceres₁ / qualia) no pueden poseer, como componentes, acaecimientos externos. Es la búsqueda de Descartes y Locke de un indicio causal, externo, en las ideas de la mente. Un residuo del marco cartesiano parece regir el uso de estas analogías, en las que, subrepticamente, el componente fenoménico reaparece una y otra vez cuando se lo desea reducir o eliminar. Este residuo es, quizás, la fuerte demarcación entre “lo interno” y “lo externo”.

Por supuesto, no todas las analogías reductivas tienen componentes fenoménicos en su analogado, como ocurre con las siguientes:

3. *“Así como el desarrollo de las ciencias biológicas ha ido ganando impulso con respecto al vitalismo [Análogo], es razonable esperar que el desarrollo de la neurobiología u otras ciencias hagan lo mismo con respecto a los impulsos antifisicalistas [Analogado]”⁷⁶*
4. *“Así como el flogisto simplemente se mostró como inexistente y por lo tanto las discusiones en torno a él se dejaron de lado gracias al avance científico [Análogo], es posible que los términos mentales (fenoménicos) corran la misma suerte [Analogado]”*

⁷⁶ Es el caso de Papineau en *Thinking about consciousness*, citado por Horst (2007, pág. 102) “because developments in the life sciences have progressively provided alternatives to the vitalist impulse, it is reasonable to expect that future work in neuroscience or other disciplines will do the same with respect to the antiphysicalist impulse”

Las analogías 3 y 4 no cometen la petición de principio que encontramos en 1 y 2, porque “vitalismo” y “flogisto” no son conceptos fenoménicos: son términos teóricos propios de un modelo científico. Pero tampoco poseen componentes fenoménicos en su analogado: “neurobiología” y “términos mentales” no son fenómenos mentales. En definitiva, las analogías 3 y 4 no están buscando relaciones entre la fenomenología y los acaecimientos, sino relaciones entre ciencias (analogía 3) y lenguaje (analogía 4).

Podemos ver que en las analogías QED se apunta a la postulación de un modelo de lo mental en el cual la mente (analogado) es semejante a otra cosa (análogo) que de todos modos presupone a la mente. La comparación y la semejanza no se buscan, entonces, en dos ámbitos diferentes sino en el mismo, y en ese sentido no se está mostrando que el modelo de lo mental deberá seguir las pautas de un modelo físico o perceptivo, sino que en ambos casos aparecen modelos que implican lo mental.

Pero es importante notar que todos los modelos, incluso los que se proponen como modelos alternativos o “adecuados”, son ellos mismos productos de un trabajo mental colectivo (aun cuando los modelos no sean, estrictamente hablando, “mentales”, sino que pertenezcan al tercer mundo). “Scientific models are a special case of mental models generally.” (Horst, 2007, pág. 172); “[I]t is a general design principle of the cognitive architecture of humans (and other animals) that the mind possesses multiple models for understanding and interacting practically with different aspects of the world” (Pág. 174): esta afirmación nos lleva a concluir que cualquier modelo (del mundo y de la mente) ha de ser necesariamente parcial, y de ese modo un modelo de la mente no puede ser nunca una *explicación suficiente* de lo mental: parece necesario que en la explicación de lo mental (como en cualquier otro dominio), contribuyan múltiples modelos que no podrán ulteriormente unificarse. Quizás exista algún modelo unificado que explique completamente la naturaleza de lo mental, pero según la línea de razonamiento de Horst dicho modelo no es asequible a la mente humana

Cabalgando por encima de las analogías fallidas: *Enactividad* de la conciencia o *exoconciencia*.

Si nos quedáramos con la conclusión del apartado anterior, no parece haber posibilidad de lograr una explicación satisfactoria y unitaria de la conciencia,.

Horst proponía que la *unificación* de las percepciones en una experiencia (el color, el tamaño, la forma, las cuales se representan en el cerebro en módulos separados) no necesitaba de una *ulterior representación unificadora* (i.e., un “módulo unificador”). La unificación no es representable *qua* unificación; surge de una sucesión de representaciones intermodales. Los modelos mismos son *representaciones* abstractas e idealizadas. Frege señalaba que *no todo contenido de conciencia es representación y que, por lo tanto, el intelecto no solo puede contemplar lo que es contenido de conciencia* (Frege, 1974, pág. 153). Podríamos extraer algunas consecuencias de esta afirmación. Al afirmar que *no todo en la conciencia es contenido*, Frege parecía decir que *no todo contenido de conciencia está “en” la interioridad de la conciencia*, - si entendemos a la conciencia como un centro inaccesible, inmaterial y en cierto modo misterioso de nosotros mismos. *De algún modo, lo que yo capto no está en mí como sujeto “encerrado” en la interioridad de mi propia subjetividad, sino que está en “sí mismo”, en una externidad que es ajena a mi conciencia.*

Los capítulos anteriores trataron acerca de la reducción de fenómenos mentales a acaecimientos; de ideas a cualidades y de cualidades a propiedades de las cosas. Tanto en el análisis de casos de la modernidad (Descartes y Locke) como en los casos de filósofos contemporáneos (Dennett, Foss, Pautz, Matthen) el problema central con respecto a la conciencia es el de la brecha explicativa entre los fenómenos y los acaecimientos. Si hay una brecha explicativa, nuestra conciencia *siempre* está del lado “inexplicado” de la brecha. Nuestra vida consciente e incluso la explicación de procesos no conscientes requieren de nuestra conciencia en tanto destinatarios de dichas explicaciones. Una explicación cuenta como explicación si es asequible a una

conciencia –en particular, a la conciencia humana, aunque no tiene por qué estar restringida a ella. Si se asume la existencia de un *gap* explicativo, se asume, con ello, de qué lado estamos *nosotros* en tanto seres conscientes que asumen tal *gap* explicativo. Dicho en otras palabras, la aceptación del *gap* explicativo en el problema mente – cuerpo es, a su vez, la aceptación de que la propia subjetividad consciente de quien acepta ese *gap* *está del lado inexplicado del gap*. La *comprensión consciente* del *gap* explicativo es, a su vez, la *comprensión consciente de la incomprensión de las relaciones entre esa comprensión consciente y de los hechos de bajo nivel a los que dicha comprensión se reduciría*. Pero quizás no es necesaria una comprensión de la conciencia en el sentido de la elaboración de un modelo para explicar cómo la mente adquiere contenidos, uno de cuyos contenidos es la propia conciencia. Quizás hay algo en el propio concepto de “modelo representacional de la conciencia” que está mal enfocado.

Uno de los problemas con la conciencia es saber si se está mirando en el lugar adecuado. Si se observan mecanismos físicos, e incluso partes funcionales de esos mecanismos físicos hasta llegar a los elementos últimos de la realidad, la conciencia tal vez no aparezca entre ellos⁷⁷. Para encontrar conciencia, hay que mirar con un ojo diferente. Si se les pregunta a personas no familiarizadas con ciertos planteos filosóficos: “¿Dónde está su mente?”, suelen señalar sin demasiadas dudas su cabeza. Cuando se les llama la atención acerca de que las sensaciones suelen estar localizadas en el cuerpo, pueden llegar a decir que la conciencia está difundida en cada célula o cada nervio. Si no desean aceptar esto (porque parece necesario aceptar algún tipo de identidad entre cuerpo y conciencia) algunas personas se refugian en un protodualismo, en el cual la conciencia se identifica con el alma y adoptan algún tipo de conciencia cartesiana. La conciencia es una misteriosa interioridad a la que sólo se accede escudriñándose a sí mismas con “el ojo de la mente”. Aunque la conciencia es un hecho, no es algo que pueda señalarse como localizada en algún lugar, como lo están los colores, los árboles, las sillas y las universidades. Sin embargo, cuando aplicamos la estrategia reductiva, pretendemos decir que *esto* y *aquello* que

⁷⁷ A menos que algunas hipótesis acerca de la naturaleza de los fenómenos cuánticos se revele como crucial para este propósito, pero según algunos filósofos no hay un buen prospecto en esta estrategia. Cfr. Searle (2000), págs. 82 – 3.

señalamos son las causas proximales de la conciencia. De algún modo, la estrategia reductiva nos lleva a la búsqueda de objetos que podrían señalarse.

Pero la propia conciencia parece imposible de localizar, porque está del lado inexplicado del gap. Parece que la conciencia sólo es localizable por la propia conciencia o, en otras palabras, sólo la encontramos cuando nos miramos “con el ojo de la mente” y no con el ojo que mira cuando le señalan. ¿Qué ve el ojo de la mente? ¿Qué ve en el interior de su conciencia? ¿Qué conoce cuando, siguiendo el precepto socrático, usted se “conoce a sí mismo”? En vano buscará un misterioso “interior” inexistente. Si usted piensa que la conciencia es algo concéntrico, estará tentado a buscar su centro en *usted* mismo, entendiendo por “usted” algo que pasa puntualmente en su conciencia anclada a su cerebro o a su cuerpo. Todo lo que pasa “afuera” de usted no puede pertenecer a la conciencia. Pero si la conciencia es excéntrica, los hechos externos no sólo son externos en tanto que ocurren en el mundo, incluso lejos del propio cuerpo: los hechos externos son *parte integrante* (y fundamental) de la conciencia. La distinción “interno” y “externo” es evolutivamente relevante para establecer los límites entre el propio cuerpo y lo que no pertenece al cuerpo, pero no sirve para distinguir entre cuerpo y exterior por un lado, y yo consciente por el otro. Señalábamos, al final del apartado anterior, que las analogías QED parecían tener como base un resabio de cartesianismo; una fuerte demarcación entre el “afuera” y el “adentro” de la conciencia, como si los términos “afuera”, “adentro”, “objetivo” y “subjetivo” tuviesen un valor absoluto (precisamente, por no atender a este resabio de dualismo, es que todas ellas contenían elementos del “adentro”, de la propia conciencia, incluso en el análogo). Quizás, como dicen Thompson, Varela y Palacios:

There is nothing wrong in drawing a distinction between subjective and objective, or internal and external, relative to the framework of a given neurophysiological, psychophysical, or behavioral experiment. The problems arise, rather, when we attempt to force perception theory as a whole into some absolute, subjective/objective straitjacket derived from the empiricist tradition. (Thompson *et al.*, 1998, pág. 396)

Cuando analizamos las posiciones de Foss (§2.6), Pautz (§2.7) y Matthen (§2.8) encontramos que una explicación reductiva del fenómeno de la visión no podía dar

cuenta satisfactoria de a) la desaparición del fenómeno en favor de elementos funcionales o físicos; b) un concepto de “visión en colores” que fuera lo suficientemente abarcativo para todo fenómeno de visión en colores, independientemente de la especie. En particular, notábamos que la explicación del fenómeno de la visión mediante la teoría de los pares oponentes cometía el mismo error que las analogías fallidas sobre la conciencia expuestas más arriba: pretendían explicar lo fenoménico mediante un esquema que también es fenoménico.

Sostendremos que es quizás un error buscar los puentes entre lo mental (fenoménico) y lo físico como si se tratara de dos dominios desunidos. La conciencia es quizás una especie de *exoconciencia*⁷⁸, que está difuminada en lo que creemos es “exterior” a nosotros mismos:

Nosotros mismos, por nuestra propia naturaleza, estamos también distribuidos dinámicamente; cruzamos fronteras, desembarcamos y estamos situados en el entorno. La razón por la que no hemos sido capaces, hasta ahora, de entender la conciencia es que la hemos estado buscando en el lugar equivocado. (Noë, 2010, pág. 94)

Es posible, entonces, que la búsqueda de puentes nomológicos entre la “visión fenoménica” y las bases físicas de la visión no nos puedan explicar la totalidad *sin resto* del fenómeno. Pero ocurre que, quizás, el fenómeno a ser explicado no es un *quale* que “está en la cabeza”:

[E]s saludable recordar que la visión evolucionó en los animales no para que pudieran representarse el mundo en la cabeza, sino para permitir una vida comprometida, por ejemplo, perseguir presas y parejas, y evitar predadores y otros peligros (Noë, pág. 207)

Podemos entender a la visión (y, por extensión, a toda nuestra conciencia) como un caso de *evolución enactiva* (Thompson *et. al.*, 1998): el organismo y el entorno se codeterminan mutuamente:

It might be admitted that animal activity can perturb, select, or construct the environment, but since perception is considered to be fundamentally distinct from action, perception per se does not participate in animal – environment

⁷⁸ Sustantivo. De *exo-*: por fuera y *conciencia*. No debe entenderse en el sentido de una conciencia de origen extraterrestre (sentido que se suele utilizar), sino en el de una conciencia que no “reside” en el interior de un cuerpo tal como un objeto en una caja.

codetermination. (...) The first step for perceptual theory is to refuse to separate perception from action or, more generally, from perceptually guided activity. [The central fact is that] perception and action have evolved together – that perception is always *perceptually guided activity* (Thompson *et al.*, 1998, pág. 398)

No existe un mundo ya dado que el animal debe recoger con sus sentidos: la conducta perceptiva del animal constituye lo que es su mundo relevante, y de ese modo percibe lo que ha construido como entorno para sí mismo. Lo percibido no está “en la cabeza” del perceptor, sino en el entorno mismo:

The perceptual object, though experiential, is also ineliminably ecological, and so [...] is not simply “in the head”. [...] According to the enactive view, the contents of perceptual states are to be type-identified by way of the ecological properties perceived, and these ecological properties are to be type-identified by way of the states that perceives them. [...] To specify perceptual content for a given animal we must investigate the relevant environmental properties, and to determine the relevant environmental properties we must investigate the sensory – motor patterns of activity that constitute the animal’s perceptual states (Pág 401-2)

Si la percepción es una estructura dinámica que transcurre entre el animal y el entorno, cobra fuerza la idea de que la propia consciencia no debe buscarse en el interior de un cerebro, sino en la interacción entre la mente y el mundo:

La conciencia del mundo es algo que nosotros hacemos: la representamos, con la ayuda del mundo, en nuestras actividades dinámicas. No es algo que ocurre en nuestro interior . [...] [P]odemos decir que el propio mundo pertenece a la maquinaria de nuestra propia conciencia. (Noë, 2010, pág.s 90-1)

De acuerdo a la propuesta enactiva, no hay percepción sin un entorno y no hay entorno sin una coevolución entre el organismo perceptor y el entorno. De modo que una conciencia perceptora ha de surgir de estas interacciones entre el sujeto y su entorno. Buscar los puentes entre el aspecto fenoménico y el físico – funcional sólo tendría sentido si la conciencia estuviera en alguno de los extremos de ese puente, en alguna misteriosa interioridad.

Decir que el cerebro es consciente es cometer la falacia mereológica⁷⁹, pues no se le pueden adscribir atributos psicológicos a una entidad física. ¿Dónde está el yo consciente? En todo aquello que está percibiendo y todo lo que *le parece*₁. Si se acepta esto, puede decirse que hay genuinos estados cualitativos de conciencia, pero no de la conciencia en tanto algo que ocurre *en* el cerebro, sino de la *exoconciencia como hecho enactivo*, en tanto se trata de algo que ocurre en los espacios externos. Mis estados cualitativos son el gris del frente de la escuela que percibo a través de la ventana *en tanto es de la escuela y no de mi “mente”*. Esta idea no está muy alejada del realismo proyectivista de Frege: incluso la idea de que el intelecto capta entidades objetivas debería poder entenderse como el intelecto proyectado a la realidad de la que forma parte. La distinción “afuera” y “adentro” se disuelve. Los fenómenos mentales existen: son los colores y sonidos *del mundo*, los sabores de los alimentos, las texturas de las cosas. Son las sensaciones viscerales de nuestro propio cuerpo, en tanto nuestro propio cuerpo, nuestro sistema perceptivo y nuestra propia mente son partes del mundo. Como Gibbons señalaba, el parecer₁ del color rojo más cercano está a una cierta distancia de nosotros (no señalamos el rojo de un sillón como un suceso mental, sino como un acaecimiento que ocurre *fuera* de la propia mente). Podría pensarse que la conciencia está proyectada en esa distancia, no “encerrada aquí dentro” del cerebro. En ese sentido, deja de ser importante la brecha que Locke encuentra entre las cualidades secundarias y las ideas; no es necesario apelar a la arbitrariedad divina, porque el rojo es *lo que hay ahí*, lo que encontramos nosotros porque “somos nosotros ahí”. Lo mismo ocurre con las cualidades primarias: la tensión que muestra Locke con respecto a la noción de *cualidad* también puede disolverse mediante la afirmación de la exoconciencia. En la primera sección evaluamos la posibilidad de que para Locke la semejanza entre las cualidades primarias y las cosas sea literal, pero no estricta. Pero la aceptación de esta semejanza literal implica la aceptación de que en cierto modo hay una representación interna que es semejante aunque no estrictamente idéntica a la cualidad primaria instanciada. ¿Es necesario

⁷⁹ “Llamaremos «falacia mereológica» en neurociencia al error de los neurocientíficos de atribuir a las *partes* constituyentes de un animal atributos lógicamente aplicables sólo al animal como un *todo*. Llamaremos «principio mereológico» en neurociencia al principio de que los predicados psicológicos aplicables únicamente a un ser humano (u otro animal) en su totalidad no se pueden aplicar de modo inteligible a sus partes, por ejemplo al cerebro” (Bennett, M *et al*, 2008, pág. 38)

aceptar esta reduplicación: cualidad, representación literal no estricta, estado cualitativo de conciencia? Se ha interpretado a Locke como diciendo que las cualidades primarias son indiscernibles del objeto mismo. ¿Por qué no dar un paso más (uno que Locke estuvo a punto de dar) y decir que las *ideas* (y no sólo las *cualidades correspondientes*) *de forma, figura, posición y solidez* no pueden separarse del objeto? ¿Podrá ser esa la fenomenología objetiva que pedía Nagel, la “Subjetividad objetiva” con la que especulaba Gibbons o, incluso, la “fenomenología real” que parodiaba Dennett? El mundo es parte de la realidad; el cerebro es una parte del mundo; la conciencia depende del cerebro, la conciencia también debe ser parte del mundo, pero no se puede integrar a él por sus propiedades fenoménicas. ¿Y si las propiedades fenoménicas *son* el mundo? La pregunta anterior propone la fusión del fenomenalismo con el realismo (al modo de Wittgenstein): el fenomenalismo, llevado hasta sus últimas consecuencias, coincide con el realismo⁸⁰. Quizás las cualidades secundarias y primarias es todo lo que se necesita y no hace falta apelar a las ideas o a los qualia en tanto “entidades” que están “en el interior” de la mente. El problema de la distinción entre las cualidades e ideas está en la búsqueda de una contrapartida de las cualidades *en* las ideas. Pero quizás no haya contrapartida; quizás no sea -como pensó Berkeley-, que las cualidades secundarias se identifican con las ideas en tanto ambas están en la mente: hay identidad entre cualidades secundarias e ideas, pero en el sentido inverso a como las pensó Berkeley: ambas existen, pero *en una totalidad dinámica de conciencia que es excéntrica, que no puede decirse que esté “afuera” o “adentro” de los propios procesos conscientes*. Frege dio un paso decisivo en este camino, al decir que el lenguaje se refiere a algo *como estando afuera*.

Supóngase que descubrimos los circuitos neurales que provocan la conciencia. Mediante un escaneo cerebral se me muestra en vivo los circuitos de mi cerebro que son responsables de mi conciencia: los veo y digo “Así que eso que veo provoca mi conciencia”. “Eso que veo” es lo que me permite ver eso y decir “así que esto que veo provoca mi conciencia”. ¿Sería correcto que, en ese caso, yo dijera “Ahí está mi

⁸⁰ § 5.63: “Yo soy mi mundo (El microcosmos)”; § 5.64: “Se ve aquí cómo, llevado a sus últimas consecuencias, el solipsismo coincide con el puro realismo. El yo del solipsismo se contrae hasta convertirse en un punto inextenso y queda la realidad con él coordinada.” (Wittgenstein, Tractatus Logico Philosophicus, 1999)

conciencia”, “Eso soy yo”? En algún sentido es correcto, porque todo lo que percibimos es parte de nuestra conciencia, y en algún sentido la conciencia está allí, donde estoy mirando. En otro sentido no es correcto: mi conciencia no está *en un lugar* determinado; la conciencia está difundida en *todo* lo que percibo. ¿Se puede decir que en mi cerebro está, causalmente, mi conciencia? Sólo de un modo restringido: allí hay algo como una “conciencia originaria”, una disposición *a priori* a ser conciencia, pero esa disposición (esos patrones causales, por llamarlos de alguna manera) no son todavía condición suficiente para la conciencia. La conciencia sólo puede ser tal si hay contacto con el medio ambiente (naturaleza, sociedad, lenguaje), y en ese sentido todo el universo (e incluso más allá de él) es el “recinto” de la conciencia. Quizás sólo entendida de esa manera la conciencia puede corregirse la aparente paradoja que encontramos en Wittgenstein, de que el ojo no puede verse a sí mismo. ¡Pero en un sentido importante, el ojo puede verse a sí mismo! Los sentidos no funcionan gracias a la conciencia: ellos, junto con sus objetos correspondientes, *originan* la conciencia. En un sentido estricto, no hay *contenidos* de conciencia, si por “contenido de conciencia” queremos referirnos a alguna misteriosa interioridad separada del mundo: allí donde hay conciencia, hay fenómenos, pero esos fenómenos no están “dentro” de una oculta facultad “en el interior de nosotros mismos”, sino difundidos a través del mundo. No hay riesgo de caer en un pampsiquismo, pues no decimos que las paredes, el techo, las mesas y la silla que observamos son *conscientes*, sino que son *partes de una conciencia*, la del sujeto que las observa.

No podemos lograr un modelo unificado de la conciencia porque, en rigor, quizás no debamos elaborar un modelo que dé cuenta de modo unificado de dos hechos que están desunidos: en realidad, no hay dos hechos (el fenoménico y el físico), sino uno: la conciencia en el mundo.

Hemos llegado a las siguientes conclusiones:

1. Las brechas explicativas están generalizadas en la ciencia y no hay prospecto de que pueda lograrse una unidad de todos los modelos explicativos. Esta aparente aunque sugestiva imposibilidad de unificación de modelos parece residir en nuestra propia arquitectura cerebral. De todos modos, el conocimiento actual de la arquitectura

mental en conexión con la tesis de que no existe unidad en los modelos explicativos de la ciencia parece revelarnos que no podemos construir *un modelo unificado* de nuestra mente.

2. Las conclusiones de 1) sólo hacen mella en nuestra concepción de lo mental si mantenemos una separación tajante entre la “interioridad” de lo mental y la “exterioridad” de los hechos físicos o funcionales que le sirven de base. Pero puede entenderse a la conciencia como *exoconciencia*, como una entidad difundida a través de los objetos percibidos, que está presente *en todo aquello de lo que* somos conscientes. Los contenidos de conciencia no están “dentro de” una conciencia: la conciencia es literalmente (y no metafóricamente) el hecho percibido en tanto ese hecho está “afuera”. La conciencia está tan “afuera” como el color rojo que percibo en el sillón que estoy viendo. En realidad, la distinción entre “afuera” y “adentro” es una mala guía. (Que, de manera correlativa a mi percepción, “pase algo” en mi cerebro, no forma parte de mi conciencia, ni aporta indicios acerca de la “sede” del hecho consciente).

3. La conciencia es *real e irreductible*. Pero no es irreductible porque todavía no se hayan encontrado los puentes nomológicos entre el aspecto fenoménico y el aspecto físico – funcional: es irreductible porque los elementos que se buscan reducir no se encuentran anclados en una interioridad misteriosa. Un fenómeno de conciencia no es una función de un quale y un suceso neuronal: el propio fenómeno de conciencia está engarzado en el mundo.

4. El cerebro no es consciente ni es la *sede* de la conciencia: es una caña de pescar que, arrojada convenientemente, “pescar” conciencia a través de los sentidos. Pero la conciencia no es una función del cerebro: el cerebro es una de las múltiples condiciones necesarias para la conciencia, pero no hay identidad entre esta última y el cerebro.

Referencias.

- Adams, E. M. (29 de Julio de 1948). Primary and Secondary Qualities. (I. Journal of Philosophy, Ed.) *The Journal of Philosophy*, Vol. 45(16), 435 - 442.
- Alex Byrne & Michael Tye. (2006). Qualia ain't in the head. *Nous*, 40, 241 - 255.
- Arias Domínguez, A. (Enero de 2011). ¿Ha logrado Dennett quinear los qualia? Una revisión naturalista. *A Parte Rei*(73), 1-25.
- Armstrong, D. (1965). *Berkeley's philosophical writings*. Londres: Collier - McMillan.
- Averill, E. W. (Julio de 1982). The Primary -Secondary Quality Distinction. (D. U. Press, Ed.) *The Philosophical Review*, 91(3), 343 - 361.
- Belot, G., & Earman, J. (1997). Chaos out of Order: Quantum Mechanics, The Correspondence Principle and Chaos. *Studies in the History and Philosophy of Modern Physics.*, 2, 147 - 182.
- Bennett, J. (1988). *Locke, Berkeley, Hume: Temas Centrales*. México: UNAM.
- Bennett, Maxwell; Dennett, Daniel; Hacker, Peter; Searle, John. (2008). *La Naturaleza de la conciencia. Cerebro, mente y lenguaje*. Barcelona: Paidós.
- Berkeley, G. ((1710) 1985). *Principios del conocimiento humano*. Madrid: SARPE.
- Block, N. (1990). Inverted Earth. *Philosophical Perspectives*, 4, 53 - 79.
- Block, N. (2003). Mental Paint. En M. H. Rambert, *Reflections and Replies: Essays on the Philosophy of Tyler Burge*. MIT Press.
- Block, N. (s.f.). Mental Paint.
- Borges, J. L. (1983). Funes el memorioso. En J. L. Borges, *Ficciones* (págs. 109 - 119). Buenos Aires: Emecé.
- Borges, J. L. (1999). Tlön, Uqbar, Orbis Tertius. En J. L. Borges, *Ficciones* (págs. 13 - 34). Madrid: Alianza Colección.
- Cartwright, N. (1983). *How the Laws of Physics Lie*. New York: Oxford University Press.
- Cepedal, F. (1976). *Diccionario de filosofía contemporánea*. Salamanca: Sígueme.
- Cepedal, F. (1976). *Diccionario de Filosofía Contemporánea*. Salamanca: Sígueme.
- Chalmers, D. (1999). *La Mente Consciente. En busca de una teoría fundamental*. Barcelona: Gedisa.
- Chisholm, R. M. (1982). *Teoría del Conocimiento*. Madrid: Tecnos.

- Davidson, D. (1970). Mental Events. *Experience and Theory*(44), 43 - 64.
- De Valois, R. L; Abramov, I; Jacobs, G. H. (1967). Analysis and coding of the response patterns of LGN cells. *Journal of the Optical Society of America*(56), 966 - 977.
- Dennett, D. ((privately printed, available from the American Philosophical Association. 1987)). *The Philosophical Lexicon* (8 ed.).
- Dennett, D. (1987). *The Intentional Stance*. England: The MIT Press.
- Dennett, D. (1988). Quining Qualia. En A. & Marcel, *Consciousness in Modern Science*. Oxford University Press.
- Dennett, D. (1995). *La Conciencia Explicada*. Barcelona: Paidós.
- Dennett, Daniel; Hofstadter, Douglas. (1983). *El ojo de la mente. Fantasías y reflexiones sobre el yo y el alma*. Buenos Aires: Sudamericana.
- Descartes, R. (1951). *Los Principios de la Filosofía*. Buenos Aires: Losada.
- Descartes, R. (1957). Meditaciones Metafísicas. En R. Descartes, *Obras escogidas*. Buenos Aires: Sudamericana.
- Descartes, R. (1967). *Obras Escogidas*. Buenos Aires: Sudamericana.
- Descartes, R. (1981). Dióptrica. En R. Descartes, *Discurso del Método, Dióptrica, Meteoros y Geometría*. Madrid: Alfaguara.
- Descartes, R. (s.f.). *Meditaciones Metafísicas*.
- Downing, L. (1998). The Status of Mechanism in Locke's Essay. *The Philosophical Review*, 381 - 414.
- Dretske, F. I. (1995). *Naturalizing the Mind*. Cambridge: MIT Press.
- Eccles, John & Zeier, Hans. (1985). *El Cerebro y la Mente. Reflexiones biológicas sobre la prehistoria, naturaleza y porvenir del hombre*. Barcelona: Herder.
- Eddington, A. S. (1937). *La Naturaleza del Mundo Físico*.
- Felleman, D. J., & van Essen, D. C. (1991). Distributed Hierarchical Processing in the Primate Visual Cortex. *Cerebral Cortex*(1), 1 - 47.
- Fodor, J. (1983). *Modularity of the Mind*. MA: Mit Press.
- Foss, J. (2009). The Scientific Explanation of Colour Qualia. *Dialogue*(48), 479 - 512.
- Frege, G. (1974). El Pensamiento, Una Investigación Lógica. En G. Frege, *Escritos Lógico - Semánticos* (págs. 136 - 158). Madrid: Tecnos.
- Frege, G. (1974). Función y Concepto. En G. Frege, *Escritos Lógico - Semánticos* (págs. 11 - 31). Madrid: Tecnos.

- Frege, G. (1985). Sobre Sentido y Referencia. En G. Frege, *Estudios sobre Semántica* (U. Moulines, Trad.). Barcelona: Ariel.
- García Carpintero, M. (1996). *Las palabras, las ideas y las cosas*. Barcelona: Ariel.
- Gibbons, J. (Diciembre de 2005). Qualia: They're not what they seem. (Springer, Ed.) *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 397 - 428.
- Gould, S. J., & Lewontin, R. C. (1979). The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptational Programme. *Proceedings of the Royal Society of London, B* 205, 581 - 598.
- Gregory, R. L. (1968). Perceptual illusions and brain models. *Proc. Royal Society(B* 171), 279 - 296.
- Hacking, I. (1979). *¿Por qué el lenguaje importa a la filosofía?* Buenos Aires: Sudamericana.
- Hardin. (1984). A New Look at Color. *American Philosophical Quarterly*(21), 125 - 133.
- Hardin. (1988). *Color for Philosophers: Unweaving the Rainbow*. Indianapolis: Hackett.
- Hellman, G. (1999). Reduction (?) of What (?): Comments on L. Sklar's "The Reduction (?) of Thermodynamics to Statistical Mechanics". *Philosophical Studies*, 1 - 2(95), 200 - 13.
- Hering, E. (1878). *Zur Lehre Vom Lichtsinne. Sechs Mittheilungen an die Kaiserliche Akademick der Wissenchaften in Wien*. Viena: Carl Gerlord's Sohn.
- Hilbert, D. (1992). What is Color Vision? *Philosophical Studies*(68), 351 - 370.
- Hobbes, T. (1979). *Leviatán*. Madrid: Editora Nacional.
- Horst, S. (2007). *Beyond Reduction. Philosophy of Mind and Post - Reductionist Philosophy of Science*. New York: Oxford University Press.
- Hume, D. ((1739) 1981). *Tratado de la naturaleza humana* (Vol. vol. 1). Madrid: Editora Nacional.
- Humphrey, N. (1995). *Una Historia de la Mente*. Barcelona: Gedisa.
- Jackson, F. (1982). Epiphenomenal Qualia. *Philosophical Quarterly*, 32, 127 - 36.
- Jacovides, M. (Octubre de 1999). Locke's Resemblance Theses. (D. U. Press, Ed.) *The Philosophical Review*, 461 - 496.
- Kim, J. (2000). *Mind in a Physical World: An Essay on the Mind- Body Problem and Mental Causation*. MIT Press.
- Kitcher, P. (1984). 1953 and All That: A Tale of Two Sciences. *Philosophical Review*(93), 335 - 373.
- Leibniz, G. W. (1714). *Monadología*.

- Levine, J. (1983). Materialism and Qualia: The Explanatory Gap. *Pacific Philosophical Quarterly*, 354 - 361.
- Lewis, D. (1997). Naming the colours. *Australasian Journal of Philosophy*(75), 325 - 42.
- Lewontin, R. C. (1983). Biological Determinism. En *Tanner Lectures on Human Values*. Salt Lake City: University of Utah Press.
- Locke, J. ((1710) 1956). *Ensayo sobre el entendimiento humano*. México - Buenos Aires: Fondo de Cultura Económica.
- Lorenzo, Á. M. (1996). La influencia de la física corpuscular en la filosofía de Locke. *Thémata*, 107 - 125.
- Matthen, M. (Enero de 1999). The Disunity of Color. (D. U. Press, Ed.) *The Philosophical Review*, 108(1), 47 - 84.
- Maxwell Bennett, Daniel Dennett, Peter Hacker, John Searle. (2008). *La Naturaleza de la Conciencia. Cerebro, Mente y Lenguaje*. Barcelona: Paidós.
- Mc Ginn, C. (1995). Consciousness Evaded: Comments on Dennett. *Philosophical Perspectives*, 9, *AI, Connectionism and Philosophical Psychology*, 241 - 250.
- Merleau Ponty, M. (1994). *Fenomenología de la Percepción*. Buenos Aires: Planeta - Agostini.
- Mollon, J. D. (1997). "...On the Basis of VELOCITY Clues Alone": Some Perceptual Themes 1946 - 1996. *The Quarterly Journal of Experimental Psychology*, 859 - 878.
- Morowitz, H. J. (1983). El Redescubrimiento de la Mente. En D. R. Daniel Dennett, *El Ojo de la Mente* (págs. 47 - 66). Buenos Aires: Sudamericana.
- Mosterín, J. (1981). *Perceptos y Conceptos*. Barcelona: Salvat.
- Nagel, E. (1974). What Is It Like To Be a Bat? *Philosophical Review*(4), 435 - 450.
- Nagel, E. (1983). ¿Cómo es ser murciélago? En D. R. Dennett, *El Ojo de la Mente* (págs. 505 - 534). Buenos Aires: Sudamericana.
- Neitz, M. & Neitz, J. (1998). Molecular Genetics and the Biological Basis of Color. En R. K. W.G. Backhaus, *Color Vision: Perspectives From Different Disciplines*. Berlín: Der Greuter.
- Noë, A. (2010). *Fuera de la Cabeza*. Barcelona: Kairós.
- Oppenheim, Paul; Putnam, Hilary. (1958). Unity of Science as a Working Hypothesis. En H. Feigl, M. Scriven, & G. Maxwell, *Concepts, Theories and the Mind- Body Problem*. Minneapolis: University of Minesota Press.
- Papineau, D. (2002). *Thinking About Consciousness*. Oxford: Oxford University Press.
- Pautz, A. (2006). Sensory Awareness Is Not a Wide Physical Relation: An Empirical ARgument Against Externalist Intentionalism. *Nous*(40:2), 205-240.

- Pautz, A. (2010). Do Theories of Consciousness Rest on a Mistake? *Philosophical Issues*, 20 , 333 - 367.
- Pereda, C. (1994). *Razón e Incertidumbre*. México: Siglo XXI.
- Peters, R. (1956). *Hobbes*. Penguin: Harmondsworth.
- Platón. (2009). *Cratilo o de la exactitud de los nombres*. (E. d. Arcis, Ed.) Obtenido de <http://www.philosophia.cl/biblioteca/platon/Cratilo.pdf>
- Popper, K. (1986). *El universo abierto*. Madrid: Tecnos.
- Priest, G. (1989). Primary Qualities are Secondary qualities Too. *British Journal of Philosophy of science*(40), 29-37.
- Primas, H. (1983). *Chemistry, Quantum Mechanics and Reductionism*. Berlin: Springer - Verlag.
- Primas, H. (1991). Reductionism: Palaver Without Precedent. En E. Agazzi., *The Problem of Reductionism in Science*. Dordrecht: Kluwer.
- Primas, H. (1998). Emergence in the Exact Sciences. *Acta Polytechnica Scandinavica*(91), 83 - 98.
- Putnam, H. (1960). Minds and Machines. En S. H. ed., *Dimensions of Mind: A Symposium* (págs. 148 - 180). New York: New York University Press.
- Putnam, H. (1984). Mentes y Máquinas. En A. M. Alan Ross Anderson, *Controversia sobre Mentes y Máquinas* (págs. 113 - 149). Barcelona: Tusquets.
- Putnam, H. (1995). El significado de "Significado". En L. M. (Ed.), *La búsqueda del Significado* (págs. 131 - 193). España: Tecnos - Universidad de Murcia.
- Rabossi, E. (1982). Lenguaje, Pensamiento y Realidad. Peri Hermeneias 16a. *Revista Latinoamericana de Filosofía*(8).
- Ross, P. (Octubre de 2001). Qualia and the Senses. *The Philosophical Quarterly*, 51(205), 495 - 511.
- San Martín Sala, J. (1987). *La Fenomenología de Husserl como Utopía de la Razón*. Barcelona: Anthropos Editorial.
- Searle, J. (1980). Minds, Brains and Programs. *The Behavioral and Brain Sciences*, 3, 417 - 457.
- Searle, J. (2000). *El Misterio de la Conciencia*. Barcelona: Paidós.
- Sharma, A; Angelucci, A; Sur, M. (2000). Induction of Visual Orientation Modules in Auditory Cortex. *Nature*, 404(6780), 820 - 821.
- Sklar, L. (1999). The reduction (?) of thermodynamics to Statistical Mechanics. *Philosophical Studies*, 1 - 2(95), 187 - 199.

- Smullyan, R. (1983). Pesadilla Epistemológica. En D. R. Dennett, *El Ojo de la Mente* (págs. 537 - 555). Sudamericana.
- Strawson, P. (1995). Sobre el Referir. En L. M. Valdés Villanueva, *La Búsqueda del Significado. Lecturas de Filosofía del Lenguaje*. (págs. 57 - 82). Salamanca: Tecnos - Universidad de Murcia.
- Sur, M., P. E. Garraghty, Roe, A. W. (1988). Experimentally Induced Visual Projections into Auditory Thalamus and Crotex. *Science*, 242(4884), 1437 - 1441.
- Thompson, Evan; Palacios, Adrián; Varela, Francisco J. (1998). Ways of Coloring: Comparative Color Vision as a Case Study for Cognitive Science. En F. J. Newmeyer, *Language Form and Language Function*. The MIT Press.
- Tran-Duc-Thao. (1971). *Fenomenología y Materialismo Dialéctico*. Buenos Aires: Nueva Visión.
- Tye, M. (Diciembre de 1993). Reflections on Dennett and Consciousness. (T. U. College, Ed.) *Philosophy and Phenomenological Research*, LIII(4), 893 - 898.
- Warnock, J. (1953). *Berkeley*. Penguin: Hammondsworth.
- Williams, J. C. (Octubre de 1969). Are Primary Qualities Qualities? (B. Publishing, Ed.) *The Philosophical Quarterly*, 19(77), 310 - 323.
- Wilson, M. (2006). *Wandering Significance*. Nueva York: Oxford University Press.
- Wilson, M. (2006). *Wandering Significance*. New York: Oxford University Press.
- Wittgenstein, L. (1988). *Investigaciones Filosóficas*. (A. G. Moulines, Trad.) México - Barcelona: Instituto de Investigaciones Filosóficas UNAM - Editorial Crítica.
- Wittgenstein, L. (1999). *Tractatus Logico Philosophicus*. (J. M. Reguera, Trad.) Madrid: Alianza Editorial.
- Yost, R. M. (Enero de 1951). Locke's Rejection of Hypotheses about Sub-Microscopi Events. *Journal of History of Ideas*, 12(1 (Jan. 1951)), 111 - 130.