



ISSN: 1577-2292
SLMFCE

Contenido:

Editorial	1
Primer Premio TMF SLMFCE	2
Segundo Premio TFM SLMFCE	31
Tercer Premio TFM SLMFCE	67

Editan:
Antonio
Blanco Salgueiro
Concha
Martínez Vidal

Maqueta:
Patricia García
Rodríguez

Boletín de la Sociedad de Lógica, Metodología y Filosofía de la Ciencia

Especial

Noviembre de 2017

Editorial

Estimados/as socios/as:

Este boletín incluye los tres trabajos que resultaron premiados en la segunda edición del premio al mejor trabajo fin de máster, creado por La Sociedad de Lógica, Metodología y Filosofía de la Ciencia en España en 2016. Esta edición del premio invitaba a participar a los estudiantes recién graduados de las universidades españolas que hubieran defendido sus TFM's en el curso 2015-16. Los requisitos de la convocatoria establecían que los trabajos candidatos debían versar sobre temas que estuvieran dentro del área de interés de la sociedad, haber obtenido una calificación de sobresaliente 9,5 o superior, y no haber sido publicados con anterioridad.

Se recibieron un total de 7 trabajos fin de máster de entre los cuales, y tras ser sometidos a un proceso de evaluación ciega –proceso en virtud del cual al menos dos expertos evaluaron cada uno de los trabajos presentados–, resultaron ganadores los que ahora publicamos.

1º) **Simetría y similitud: la concepción transformacional de la similitud.**

Autor: **Javier Belastegui-Lazcano**
Máster en Filosofía, Ciencia y Valores
UPV-EHU
Director: Thomas Mormann

2º) **El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas.**

Autor: **Marta Serrapio Lavandeira**
Máster Universitario en Filosofía
Teórica y Práctica-UNED
Director: Carlos Solís Santos

3º) **Inferencialismo en el discurso ético.**

Autor: **Llanos Navarro Laespada**
Máster en Lógica y Filosofía de la
Ciencia-UGR
Directora: María José Frápolli Sanz

La Junta de la SLMFCE se congratula de la respuesta recibida por la convocatoria, manifiesta su intención de continuar con esta actividad y confía en que contribuya a fomentar y dar visibilidad al trabajo realizado en el marco de los diferentes másteres que se ocupan de las temáticas dentro del área de interés de la sociedad. Finalmente, nos gustaría aprovechar esta ocasión para felicitar nuevamente a los premiados, a sus directores, y a las instituciones implicadas.

Un cordial saludo,

Concha Martínez
Presidenta de la SLMFCE

www.solofici.org



PRIMER PREMIO: Simetría y similitud: la concepción transformacional de la similitud

Javier Belastegui-Lazcano¹

Director: Thomas Mormann

Máster en Filosofía, Ciencia y Valores UPV-EHU²

INTRODUCCIÓN

La reflexión acerca de la naturaleza de la semejanza es antigua, no obstante, la aparición de modelos formales y teorías empíricas (e.g. psicológicas) sobre la misma es relativamente reciente y su impacto merece especial consideración. Por esta razón, el objetivo de este escrito es doble. Por un lado, se trata de defender un cúmulo de truísmos relativos a la naturaleza de la semejanza (la 'concepción clásica') de varias objeciones que ha recibido desde la filosofía y la psicología cognitiva. Esta concepción sostiene que la semejanza entre los objetos se debe a que comparten ciertas propiedades. Cuantas más propiedades comparten los objetos, más similares son. Por otro lado, se pretende introducir y defender una caracterización distinta de la similitud, la transformacional, que puede extraerse de varias teorías pertenecientes a la psicología y las matemáticas, y que es en cierto sentido compatible con la anterior.

Para ilustrar la concepción transformacional de la similitud me gustaría acudir brevemente al famoso cuento de Borges, "La Biblioteca de Babel". En este cuento, el narrador dice ser un habitante de una enorme biblioteca formada por pisos llenos de salas hexagonales conectadas por escaleras. Cada sala de la Biblioteca contiene cinco estanterías y cada estantería 32 libros. Cada libro tiene 410 páginas, cada página 40 renglones y cada renglón 80 letras. Cada letra es uno de los 25 símbolos de un alfabeto (22 letras, espacio, coma y punto). Los libros están distribuidos aleatoriamente por la biblioteca, y ésta contiene todos los libros de las mencionadas características que podrían obtenerse combinando los 25 símbolos de diferentes maneras a lo largo de sus 410 páginas. Así que, a pesar de que la Biblioteca es gigantesca, el número de libros que hay en ella es estrictamente finito. Los eruditos de la Biblioteca dedican sus vidas enteras (sin mucho éxito) a buscar libros cuyas páginas contengan alguna combinación de símbolos que tenga sentido en alguna lengua conocida. Algunos esperan encontrar libros proféticos que les revelen sus últimos días o compendios perfectos sobre la naturaleza última del mundo.

Mi pregunta es la siguiente. En esta caótica biblioteca, ¿podríamos proporcionar a sus desesperados habitantes alguna manera útil de (re)ordenar los libros de modo que eso les permitiera volver a encontrarlos cuando quisieran? Los libros

"Cuando algunos objetos son semejantes entre sí la semejanza salta a la vista desde el principio -o salta, más bien, a la mente- y raramente hará falta otro examen"
(Hume, Tratado, libro I parte 3, sección I)

tienen títulos, pero como señala el narrador del relato, los títulos no indican nada sobre lo que está escrito en los libros. Así que ordenar lexicográficamente los libros por sus títulos no sería una buena forma de organizarlos. Pero hay otra forma de hacerlo, recordemos que:

"cada ejemplar es único, irremplazable, pero (como la Biblioteca es total) hay siempre varios centenares de miles de facsímiles imperfectos: de obras que no difieren sino por una letra o por una coma." (Borges, 1997 [p.95]).

Es decir, para cada libro de la biblioteca podríamos encontrar (en principio) todos los libros resultantes de cambiar exactamente uno de sus símbolos por otro. Estos libros serían, dice Borges, *facsímiles imperfectos* del primer libro, serían los libros más similares al libro mencionado. Después podríamos tomar aquellos libros que se diferenciaran del primer libro en exactamente dos símbolos. Y podríamos continuar así hasta ordenar todos los libros de acuerdo a su semejanza respecto al libro de referencia. ¿Se debe esta similitud a las propiedades compartidas entre los libros? En cierto sentido sí, al menos a las propiedades compartidas por las partes de los libros. Si consideráramos toda la estructura interna de un libro (la ordenación de los símbolos en determinados renglones y la de los renglones en páginas) como un cúmulo de (parte de) sus propiedades (o de propiedades de sus partes), comparar dos libros como lo hemos hecho sería equivalente a buscar coincidencias o divergencias en dicha estructura interna. Pero en esta explicación nos estaríamos dejando algo. Lo que falta es el *criterio* que nos ha permitido comparar los libros, y que consiste en tomar a un libro como referente y realizar cambios o transformaciones cada vez más complejas, mayores o 'permisivas' en él. O si se prefiere, tomar un libro aleatoriamente y contar el número de cambios necesarios para convertirlo en el referente. Son estos cambios o transformaciones los que nos han permitido ordenar 'fácilmente' (al menos teóricamente) los libros por su similitud respecto al original. Ese criterio nos ha permitido además *clasificar* los libros. Todos los libros que consisten en una variación de una letra respecto al original pertenecen a una misma clase, todos los que consisten en una variación de dos letras pertenecen a una misma clase, etc. Estas clases están, a su vez, ordenadas por la complejidad o generalidad del cambio que se ha de realizar en el libro original para obtenerlas. Si tuviéramos que colocarlos de nuevo en las estanterías, los libros más próximos al original serían aquellos que contuvieran el menor número de cambios. Los más lejanos serían aquellos que fueran el resultado de un mayor número de transformaciones en el original. Así podríamos, además de ordenar los libros espacialmente, agruparlos por bloques.

1. Actualmente soy Investigador Predoctoral en formación en el Grupo de Investigación PRAXIS del Departamento de Lógica y Filosofía de la Ciencia de la Facultad de Educación, Filosofía y Antropología (HEFA) de la Universidad del País Vasco/Euskal Herriko Unibertsitatea (UPV/EHU), y beneficiario del Programa Predoctoral de Formación de Personal Investigador No Doctor del Departamento de Educación, Política Lingüística y Cultura del Gobierno Vasco.

2. El presente trabajo se elaboró gracias a la concesión de una beca de la UPV/EHU para la realización de Másteres Universitarios durante el curso académico 2015-2016.

PRIMER PREMIO: Simetría y similitud: la concepción transformacional de la similitud

La concepción transformacional de la similitud que voy a analizar en este escrito propone algo parecido a lo que hemos hecho con los libros de la Biblioteca, pero con *cualquier* objeto. La estructura de este escrito es la siguiente. Primero se presenta la concepción clásica de la similitud y algunas críticas que le han sido formuladas tanto desde la filosofía como desde la psicología cognitiva. Se responde a dichas críticas. En segundo lugar, se presentan varios ejemplos de teorías pertenecientes a la psicología cognitiva y a la geometría que manejan una noción transformacional de la similitud. En tercer lugar se presenta el núcleo de la concepción transformacional de la similitud y se ofrece un pequeño análisis formal de la misma. Se concluye con unos comentarios de carácter general.

I. LA CONCEPCIÓN CLÁSICA DE LA SIMILITUD



I.1. Formulación de la concepción clásica de la similitud

La “concepción clásica de la similitud” es un nombre para un cúmulo de truismos acerca de la relación de similitud (absoluta y comparativa). La concepción clásica de la similitud sostiene (por lo menos) las dos siguientes tesis (e.g. (Carnap, 1928)):

- (1) La similitud (absoluta) es una relación binaria, reflexiva y simétrica.
- (2) Dos objetos X e Y son similares si tienen alguna propiedad en común.

De acuerdo con (1) todo objeto es similar a sí mismo, y si un objeto se parece a otro entonces el segundo se parece al primero. Es importante observar que la similitud no es necesariamente transitiva. Un ejemplo típico de la no-transitividad de la similitud (ya detectado por H. Poincaré) es el de la gradación de los colores. En la siguiente sucesión, cada par de colores adyacentes son similares entre sí, pero el primero no es similar al último³:



Figura 1. No-transitividad de la similitud.

La condición (2) establece una relación entre la similitud de los objetos y las propiedades compartidas por éstos. La interpretación adecuada de (2) es lo que divide a los metafísicos en realistas y nominalistas. Para los Realistas (2) es una definición de la similitud en términos de la coincidencia en algún universal. Es decir, dos objetos son similares porque instancian un mismo universal (sea una propiedad o una relación), esté éste necesariamente presente o no en el propio objeto.

3. Esta figura y todas las que siguen se han realizado con el programa GeoGebra, de uso gratuito.

Para los nominalistas (2) es una condición a satisfacer por las correspondientes entidades candidatas (conjuntos –de posibilidad-, sumas mereológicas, círculos de similitud, predicados, conceptos) a cumplir el rol de propiedades en lugar de los universales, tanto si asumen como si no que la similitud es primitiva e irreducible⁴.

El contraejemplo que hemos invocado contra la necesidad de la transitividad de la similitud es también problemático para (2). Tomemos un par de colores cualesquiera de la sucesión. Son similares, pero ¿qué propiedad tienen en común? Desde luego no tienen el mismo tono, porque son tonos estrictamente diferentes. Una manera típica de escapar del problema es reformulando (2) como (2*) (e.g. Armstrong, 1978):

- (2*) Dos objetos X e Y son similares si tienen alguna propiedad en común o tienen propiedades similares.

Obviamente, (2*) exige caracterizar la similitud entre propiedades de alguna otra manera a como lo hace (2), para evitar la circularidad. Para eso se puede suponer que las propiedades son similares si, a su vez, comparten ciertas propiedades de un grado superior, o si son diferentes propiedades determinadas de una misma propiedad determinable^{5 6}.

A menudo cuando se habla de ‘propiedades’ se pretende hacer referencia también a las relaciones. Así que podemos ampliar (2*) así:

- (2**) Dos objetos X e Y son similares si o bien tienen alguna propiedad común, o bien tienen propiedades similares, o bien están en una misma relación con otros objetos, o bien están en relaciones similares con otros objetos.



En cualquier caso, la concepción clásica atribuye las propiedades lógicas descritas en (1) a la relación de similitud y explica la similitud entre objetos como relativa a las propiedades (incluyendo posiblemente las relaciones) que comparten tales objetos. A estas dos tesis se les suelen añadir otras como las siguientes:

5. La diferencia entre (2) y (2*) está ya en (Carnap, 1928), es la diferencia entre la similitud como identidad parcial y la similitud como similitud parcial.

6. Es bien sabido que esto todavía no resuelve el problema. Por ejemplo, dos colores determinados pertenecientes a un mismo determinable pueden ser más similares entre sí que respecto a un tercer determinado del mismo determinable. No sirve de nada apelar a las propiedades (e.g. psicológicas) de los colores (tono, saturación o brillo), porque esos determinados no tienen la misma propiedad sino propiedades parecidas. Esto se ve fácilmente en el modelo espacial del color: dos puntos, como n-tuplas, pueden tener componentes próximos pero distintos para una misma coordenada.

PRIMER PREMIO: Simetría y similitud: la concepción transformacional de la similitud

(3) Un objeto Y se parece más a otro objeto X que un objeto Z si X e Y comparten más propiedades que las que comparten X e Z⁷.

(3) establece que la similitud entre los objetos crece con el número de propiedades compartidas, es una forma típica de caracterizar una similitud comparativa triádica (e.g. Armstrong, 1978). Para algunos, como (Quine, 1970), la similitud comparativa triádica se comporta mejor que la diádica (que es según Quine un tanto 'defectuosa').

Por último, la semejanza entre los objetos permite realizar clasificaciones, dado que los objetos que son de un mismo tipo o pertenecen a una misma clase coinciden en ciertas propiedades (e.g. natural):

(4) Varios objetos X_1, \dots, X_n pertenecen a una misma clase (e.g. clase natural) o son de un mismo tipo si comparten ciertas propiedades q_1, \dots, q_n .



Creo que (1)-(4) son tesis *prima facie* aceptables y ampliamente aceptadas. No obstante, en las secciones que siguen presentaré algunas de las críticas que se han lanzado contra las mismas.

2. CRÍTICAS A LA CONCEPCIÓN CLÁSICA DE LA SIMILITUD

La concepción clásica de la similitud ha recibido ataques desde diversos flancos. En esta sección me voy a centrar en algunas objeciones que provienen de la filosofía y de la psicología cognitiva.

2.1. Críticas desde la filosofía

2.1.1. Críticas de Quine

Desde la filosofía se han lanzado críticas contundentes contra la noción de similitud. En (Quine, 1970) se formulan dos objeciones principales. La primera consiste en que la relación binaria de similitud genérica tiene serios problemas lógicos, con lo cual Quine propone comenzar con una relación triádica $T(x,y,z)$ como "y se parece más a x que z":

"We cannot easily imagine a more familiar or fundamental notion than this, or a notion more ubiquitous in its applications. On this score it is like the notions of logic: like identity, negation, alternation and the rest. And yet, strangely, there is something logically repugnant about it. For we are baffled when we try to relate the general notion of similarity significantly to logical terms." (Quine, 1970)

7. Obsérvese que (3) se formula en términos del número de propiedades compartidas por pares de objetos. Esta es una formulación típica, pero no es lo suficientemente explicativa. No nos interesa saber sólo cuántas propiedades comparten los objetos al compararlos, sino cuáles.

La segunda consiste en que tanto la noción de similitud como la de clase natural tienden a ser sustituidas por otras nociones más específicas y más apropiadas para las teorías de las diferentes disciplinas científicas, lo cual haría difícil (sino imposible) una teoría filosófica general que abarcara todas esas nociones más específicas:

www.solofici.org

"It is a mark of maturity of a branch of science that the notion of similarity or kind finally dissolves, so far as it is relevant to that branch of science." (Quine, 1970)⁸

Respecto a la primera de las críticas de Quine, ya hay una noción matemáticamente precisa de la similitud. La teoría de las estructuras de similitud (o "espacios de tolerancia"), que tiene su origen por un lado en (Carnap, 1923), (Carnap, 1928) y por otro lado en (Schreider, 1975), y que formaliza la similitud como una relación binaria, reflexiva y simétrica, ha sido desarrollada y aplicada. Sobre sus propiedades formales en particular, en (Mormann, 2009) se establecen conexiones estrechas entre la teoría de las estructuras de similitud y otras ramas de las matemáticas (como la teoría de retículos o la topología). En ese mismo artículo se demuestra también como definir una similitud triádica a partir de una similitud binaria primitiva. Algunas de las críticas concretas que Quine lanza a la similitud binaria proceden de (Goodman, 1953) y afectan en particular al método constitucional conocido como 'quasianálisis'. Muchas de esas críticas han sido ya ferozmente respondidas, entre otros, por (Brockhaus, 1963), (Mormann, 1994), (Mormann, 2009) y (Leitgeb, 2007). En este nuevo marco formal el método del quasianálisis se ha subsumido como un caso especial de la teoría representacional de la medida de (Suppes; Krantz; Luce; Tversky, 1989) y se puede entender como un tipo de representación estructural (Swoyer, 1991) de un dominio de objetos similares entre sí. No es tarea del presente trabajo entrar en los detalles de esas discusiones pero en cualquier caso, creo que las afirmaciones de Quine sobre la *alogicidad* de la relación de similitud binaria están sobradamente refutadas.

Por otro lado, según Quine la similitud genérica basada en la coincidencia/divergencia de propiedades entre dos objetos irá desapareciendo paulatinamente de las disciplinas científicas. Tomemos por ejemplo el caso de la biología, que el propio Quine menciona superficialmente en su artículo. Aunque Quine no trate estos problemas en su artículo, vamos a construir un Quine* que expanda esa breve

8. Aunque Quine hace la siguiente afirmación, un tanto más esperanzadora: "Perhaps the branches of sciences could be revealingly classified by looking to the relative similarity notion that is appropriate to each. Such a plan is reminiscent of Felix Klein's so-called Erlangerprogramm in geometry, which involved characterizing the various branches of geometry by what transformations were irrelevant to each. But a branch of science would only qualify for recognition and classification under such a plan when it had matured to the point of clearing up its similarity notion." (Quine, 1970).

PRIMER PREMIO: Simetría y similitud: la concepción transformacional de la similitud

crítica examinando algunos aspectos importantes de la taxonomía biológica. La similitud parece, a primera vista, una noción indispensable para la taxonomía biológica: varios individuos se agrupan como miembros de una misma especie por su semejanza, es decir, porque comparten ciertas propiedades. Del mismo modo, el resto de taxones superiores agruparían a las especies (u otros taxones) por su semejanza, por las propiedades que comparten. Pero la cuestión no es tan sencilla. En primer lugar, la tesis de que las especies son clases naturales cuyos miembros están agrupados porque poseen ciertas propiedades consideradas como necesarias y suficientes



para su pertenencia a las mismas ha sido ferozmente disputada (Diéguez, 2012). Esto va contra (4). En segundo lugar, las tres escuelas taxonómicas en disputa, los cladistas, los evolucionistas y los fenetistas, tienen opiniones muy diferentes respecto al criterio bajo el cual se realizan dichas clasificaciones (Ereshefsky, 2004). Las tres escuelas difieren en el tipo de propiedades que consideran como teóricamente relevantes. En particular, esto conlleva una restricción en (2), ya que por un lado dos especímenes podrían ser similares pero no tener propiedades *teóricamente relevantes* en común, con lo cual la similitud no sería un indicio de la presencia de propiedades comunes. Por otro lado, dos especímenes podrían tener cierta propiedad (teóricamente relevante) en común sin que dicha propiedad estuviera estrechamente relacionada con la similitud.

Los evolucionistas y los cladistas, por ejemplo, son mucho más restrictivos que los fenetistas en la selección de propiedades teóricamente importantes. Las propiedades que ellos aceptan son *homologías*, que son rasgos comunes heredados de un determinado ancestro común más próximo que otro (cladistas y evolucionistas), o relativos a haberse adaptado exitosamente al entorno (evolucionistas). Es decir, son propiedades que resultan de estar en ciertas relaciones genealógicas con otros organismos/especies, o el resultado de haber interactuado de una determinada manera con otros organismos/especies y el entorno natural.



Lo que importa a los evolucionistas y a los cladistas no son los rasgos observables de los organismos, ni su parecido estructural o funcional, ni en general ciertos cúmulos de propiedades que caractericen unívocamente a los taxones, sino las relaciones que les han llevado a tener algunas de esas propiedades. En realidad son esas relaciones históricas y ecológicas las que permiten comparar y clasificar a las especies. Los rasgos, y en particular las homologías, son sólo indicios de esas relaciones. Así, lo que distingue a unas especies de otras es su posición genealógica, y no las propiedades compartidas.

Los fenetistas, por su parte, manejan una noción de similitud más próxima a la concepción clásica. Tienen ciertas sospechas sobre que algunos rasgos hayan de ser considerados como más importantes que otros *antes* de realizar las comparaciones. Su principal objeción será, de hecho, que argumentar que hay que seleccionar ciertos rasgos como no-naturales (e.g. por no mostrar las relaciones genealógicas reales entre las especies) es circular porque para realizar una demarcación así primero se han tenido que tomar en cuenta rasgos que luego han resultado ser, desde el punto de vista del cladista o el evolucionista, 'rasgos engañosos'. Es decir, toda demarcación entre rasgos homólogos y rasgos que no lo son supone una previa comparación y clasificación que no podía recurrir a esa distinción, básicamente porque cuando uno se encuentra con especies nuevas (y no tan nuevas) que catalogar no dispone de información alguna sobre sus antepasados. Por tanto, la comparación entre especímenes por su similitud genérica (i.e. basada en rasgos no necesariamente homólogos) es epistemológicamente previa a la identificación de homologías que permitan clasificarlos de acuerdo a criterios genealógicos. Además, según los fenetistas la taxonomía debería ser una disciplina neutral e independiente de lo que el destino depare a la teoría de la evolución.

www.solofici.org

Sin embargo, las posiciones cladista y evolucionista son las posiciones taxonómicas actualmente imperantes en la biología. Esto es porque son las posiciones más compatibles con la teoría evolucionista (Ereshefsky, 2004). Si el objetivo de la teoría evolucionista es reconstruir el árbol de la vida, es decir, las relaciones filogenéticas entre las especies (y otros taxones superiores), el criterio por el cual se realizarán las clasificaciones será genealógico. Pero entonces las propiedades relevantes al comparar a unas especies con otras serán aquellas que indiquen que hay una relación genealógica interesante entre las especies. Así que se podría argumentar, con Quine*, que en la biología la noción de similitud está siendo reemplazada por otras más específicas como la de *homología*, con una gran carga teórica, y que no encajan especialmente bien con la similitud clásica. De hecho, dos especies pueden asemejarse por tener muchos rasgos o propiedades comunes que consideraríamos *naturales* (el color de los ojos, cierto patrón en el pelaje, el tamaño, la forma de una determinada parte del cuerpo, el número de dedos, el grado de concentración de una determinada sustancia, la posesión de resistencia a una bacteria, etc.) y sin embargo estar lejos la una de la otra en el árbol de la vida. Esto es porque dichos rasgos no son verdaderas homologías, no son rasgos heredados a partir de un ancestro común (más próximo que respecto de terceras especies), son sólo semejanzas funcionales o estructurales sin fundamento genealógico (o adaptativo) alguno y que se han desarrollado independientemente.

PRIMER PREMIO: Simetría y similitud: la concepción transformacional de la similitud

Por tanto, los especímenes podrían ser similares sin que hubiera una propiedad teóricamente relevante (homología) que fundamentara esa similitud, porque podrían compartir muchos rasgos homoplásicos. Por ejemplo, los delfines comunes son muy similares a los ictiosaurios, sin embargo las dos especies no están estrechamente emparentadas (Freeman, 2008). Es más, los delfines comparten más rasgos homólogos con los elefantes que con los ictiosaurios, y en ese sentido se parecen más a los primeros que a los segundos. Los rasgos que comparten los delfines y los ictiosaurios, como las aletas, la forma alargada del cuerpo o el tipo de mandíbula, son rasgos homoplásicos que resultan de su adaptación independiente a tipos de hábitat y de presas parecidos. Por otro lado, dos especímenes podrían pertenecer a especies estrechamente emparentadas genealógicamente y no parecerlo, porque la propiedad que indicara dicha relación de parentesco no fuera directamente observable. Por ejemplo, los hipopótamos comparten un ancestro común más próximo con las ballenas que con las vacas o los cerdos (Freeman, 2008). Y en ese sentido los hipopótamos se asemejan más a las ballenas que a las vacas⁹. Estos problemas también afectarían a la similitud comparativa (3), dados tres especímenes X, Y, Z, de especies diferentes, la similitud comparativa aparente podría no encajar con la determinada por las homologías. En resumidas cuentas, incluso aunque la relación de homología es algún tipo de similitud en ciertos rasgos, los rasgos homólogos son sólo indicios de la presencia de relaciones genealógicas entre las especies, que es lo que realmente determina la clasificación de las mismas. Así que el papel que cumplen en la clasificación biológica tanto las propiedades como la similitud parece satisfacer las pesimistas predicciones quineanas*.

Así que las objeciones de nuestro virtual Quine* son las siguientes:



(I) La similitud binaria genérica “x es similar a y” está siendo sustituida por la relación “x e y poseen rasgos homólogos”, o más bien, comparativamente, “x e y comparten ciertos rasgos homólogos que no comparten con z”. Esta relación es en realidad la relación

“x e y comparten un ancestro común k más próximo que el que comparte cualquiera de los dos con otro espécimen z”.

9. El parentesco entre las ballenas y los hipopótamos se ha inferido principalmente a partir de la comparación de secuencias de ADN de especímenes de sendas especies (Freeman, 2008). Así que en este caso la similitud es entre secuencias de ADN, y consiste en la coincidencia en un mayor número de posiciones de las bases nitrogenadas. O también podríamos decir: dos secuencias de ADN son más similares cuanto menor es el número de cambios que debemos hacer en (las bases de) una para obtener la otra. La comparación de secuencias de ADN como un modo de inferir la filogenia introduce más complicaciones en las objeciones que estoy considerando, así que se dejarán a un lado.

(II) La relación de homología determina ciertas propiedades o rasgos como aceptables (teóricamente relevantes) y excluye otros como inaceptables. Algunos rasgos o propiedades aparentemente ‘naturales’ pueden no ser rasgos homólogos.

(III) Dos especímenes de especies diferentes podrían ser más similares entre sí que respecto a un tercer espécimen sin compartir rasgos homólogos (e.g. delfines e ictiosaurios)¹⁰.

(IV) Dos especímenes de especies diferentes podrían compartir rasgos homólogos sin ser más similares entre sí que respecto a un tercer espécimen (e.g. hipopótamos y ballenas).

Por un lado, se puede aceptar la conclusión de Quine*(*) de que la noción de similitud es semánticamente ambigua. En ese caso tendríamos dos nociones de similitud, similitud₁ o “coincidencia en rasgos homólogos” y similitud₂ o “coincidencia en alguna propiedad”. La biología estaría sustituyendo similitud₂ por similitud₁, de acuerdo con los principios teóricos defendidos por los cladistas.

Por otro lado, se puede sostener que hay una sola noción de similitud, pero que es el contexto, los intereses teóricos u otros factores los que determinan qué tipo de propiedades se seleccionan en cada caso como el fundamento de la similitud. Esto es, la noción de similitud no sería semánticamente ambigua, sino pragmáticamente ambigua, como defiende (Niiniluoto, 1987). Puesto que la caracterización de la similitud en términos de homologías sigue satisfaciendo las condiciones (1) y (2) de la concepción clásica (aunque ahora las propiedades comunes sólo podrán ser homologías), no hay porqué suponer que se trata de otra noción de similitud diferente. Lo que ocurre es que hay principios teóricos que son los que constriñen cuáles son las propiedades que hacen que los especímenes se asemejen. Esta respuesta, que esboza Niiniluoto, me parece bastante satisfactoria y conecta con las objeciones de Goodman que se van a tratar en la sección siguiente, en particular con la objeción (4).

www.solofici.org

10. Una vez más estamos formulándolo de forma simplificada. Lo que se requiere es una versión comparativa: “Dos especímenes de especies diferentes podrían ser más similares entre sí que respecto a un tercer espécimen compartiendo un menor número de rasgos homólogos”.

PRIMER PREMIO: Simetría y similitud: la concepción transformacional de la similitud

Aunque me gustaría detenerme aquí y dar por respondidas las objeciones de Quine*, tengo que indicar que la cuestión tampoco termina de esta manera. Se puede decir que en realidad hay una tercera relación de similitud encubierta. Las homologías son rasgos compartidos por dos especies debido a la posesión de un ancestro común más próximo (que respecto a otras especies). Pero cuando se examina por primera vez especímenes cuya historia evolutiva se desconoce, no se sabe si los rasgos que se están utilizando para compararlos entre sí son homologías. Sin embargo, necesitamos encontrar rasgos homólogos para poder inferir las relaciones genealógicas entre las especies. Aquí nos enfrentamos al razonamiento circular que ya señalaron los fenetistas. La única manera de escapar del mismo es sostener que hay criterios o métodos independientes para distinguir las homologías de aquellos rasgos que no lo son. Y de hecho así es, los biólogos disponen de ciertos tests que indican si determinados rasgos son (con una mayor probabilidad que otros) rasgos homólogos. Pero estos 'tests' apelan a cierto tipo de similitud estructural (Patterson, 1982) y de similitud a lo largo de las primeras etapas del desarrollo (Ereshefsky, 2004). Esto introduce más complicaciones en el análisis, y requeriría un tratamiento exhaustivo de los métodos concretos de selección de rasgos utilizados por los biólogos, así que es más conveniente pasar a examinar lo que ocurre en otras disciplinas.

Sea cual sea el papel de la similitud en la biología, en la psicología cognitiva la noción de similitud genérica sigue siendo importante para explicar procesos cognitivos fundamentales como el aprendizaje, la categorización, la memoria, la resolución de problemas, etc. (Hahn; Chater, 1997). Por ello mismo, se dispone de varios modelos para representar la similitud entre objetos (estímulos). El modelo más popular es el de la concepción geométrica de la similitud, por ejemplo el de (Nosofsky, 1991), que representa los objetos por puntos en un espacio y la semejanza entre los objetos por la distancia entre los puntos. Es decir, cuanto más próximos están los puntos en el espacio, i.e. cuanto menor es su distancia, más parecidos son los objetos correspondientes, y cuanto más lejos están más diferentes son. Así si un punto está a distancia 0 de otro, entonces los dos objetos correspondientes son lo más parecidos posible. Por tanto, la concepción geométrica formaliza la similitud a través de la noción de espacio métrico:

Definición (Espacio métrico).



Sea M un conjunto e $d: M \times M \rightarrow \mathbf{R}^+$. Entonces (M, d) es un espacio métrico si cumple:

- (i) $d(x,y)=0$ & $d(x,y)=0$ sii $x=y$ [Minimalidad]
- (ii) $d(x,y)=d(y,x)$ [Simetría]
- (iii) $d(x,z) \leq d(x,y)+d(y,z)$ [Desigualdad Triangular]

En términos de la similitud, la minimalidad exige que todo objeto sea máximamente similar a sí mismo, e introduce una versión de la identidad de los indiscernibles de modo que si

11. \mathbf{R} es el conjunto de los números reales.

dos objetos son máximamente similares entre sí, entonces se trata del mismo objeto. La simetría supone que un objeto x es tan similar a otro y como este otro y lo es a x . La desigualdad triangular conlleva que la similitud entre dos objetos x y z siempre es mayor (o igual a) la combinación de la similitud entre x e otro y y junto con la similitud entre ese y e z ¹².

Las propiedades de los objetos están representadas por las dimensiones del espacio. Es decir, si x es un punto que representa a un objeto, sus valores (x_1, \dots, x_n) corresponden a los valores que tiene x para cada una de sus propiedades. Por ejemplo, si la dimensión 1 representa la longitud, entonces x_1 representa la longitud del objeto x . De este modo, la distancia entre los objetos tiene en cuenta las similitudes respecto a cada propiedad y produce una similitud genérica, que difiere en función de la métrica que se escoge. La coincidencia de las n -tuplas en una coordenada corresponde a la posesión de un rasgo común, i.e. $x_i=y_i$ para algún i . Nótese que incluso aunque dos objetos no tengan ningún rasgo común todavía se puede decir que sus rasgos se parecen lo suficiente, si esas coordenadas están lo suficientemente cerca. Por tanto, tiene sentido decir que los dos objetos se parecen porque tienen propiedades parecidas (y no sólo comunes).

Podría decirse que la concepción geométrica de la similitud es algo así como la versión cuantitativa del mismo concepto utilizado por la concepción clásica, que es cualitativa. Sean $x, y \in M$, sea $\varepsilon \in \mathbf{R}^+ \cup \{0\}$ fijo. Entonces definamos $x \sim_\varepsilon y$ si $d(x,y) \leq \varepsilon$. Se comprueba que \sim_ε es una similitud. Por tanto, hay una familia de similitudes que corresponden a las diferentes distancias que fijemos como determinantes de los conjuntos de objetos que son lo suficientemente parecidos entre sí. Se puede ver por tanto que hay una estrecha conexión entre el modelo clásico de la similitud y el modelo geométrico.



12. El modelo de (Nosofsky, 1991), al igual que muchos otros, toma como espacio (\mathbf{R}^n, d) donde d es la distancia de Minkowski $d(x,y) = (\sum |x_i - y_i|^p)^{1/p}$ e $x=(x_1, \dots, x_n)$ e $y=(y_1, \dots, y_n)$. Para $p=1$ se obtiene la distancia de Manhattan o 'block city', y para $p=2$ se obtiene la distancia euclídea clásica. Pero el escéptico bien podría preguntar "¿por qué estas métricas y no otras?". Ésta fue una objeción que se lanzó contra los taxónomos fenetistas: si para realizar una clasificación hay que escoger una métrica y un algoritmo de clasificación (clustering) concretos, pero el objetivo principal de los fenetistas es obtener clasificaciones objetivas que sean independientes de principios biológicos teóricos (i.e. la teoría de la evolución) que puedan sesgar la ponderación de unos rasgos (e.g. homologías) sobre otros, entonces, puesto que hay varias métricas y algoritmos de clustering disponibles, ¿no es totalmente arbitraria la selección de una determinada noción de distancia y no otra? Pero si la noción de distancia representa la (dis)similitud entre los objetos, y toda distancia pone en relación de algún modo distinto las dimensiones del espacio, i.e. las propiedades de los objetos a comparar, ¿no es arbitraria la (dis)similitud escogida entre los objetos? ¿Por qué una determinada relación entre los valores de cada una de las propiedades (e.g. euclídea) y no otra? ¿Cómo puede entonces ser objetiva la clasificación fenetista?

PRIMER PREMIO: Simetría y similitud: la concepción transformacional de la similitud

El modelo geométrico de la similitud no se utiliza sólo en psicología cognitiva, también se utiliza en las ciencias del color, la taxonomía fenetista, la estadística, la teoría de la computación (el reconocimiento de patrones), la lingüística, la teoría de la comunicación, etc. (Niiniluoto, 1987). La escuela fenetista de taxonomía numérica introdujo el modelo geométrico de la similitud en biología con la esperanza de que las clasificaciones resultantes tuvieran un mayor rigor y fueran más parecidas a las mediciones realizadas en otras disciplinas científicas (Ereshefsky, 2004). En este caso, las UTO-s (Unidades Taxonómicas Operacionales), que son los especímenes, estaban representadas por puntos y sus rasgos estaban representados por las dimensiones del espacio. La similitud estaba representada por alguna (de las múltiples) noción de distancia. La clasificación era el resultado de un cálculo de las distancias entre las UTO-s y la posterior aplicación de un algoritmo que agrupaba por clases disjuntas a los puntos más próximos entre sí (i.e. las especies más parecidas). A pesar de que, como decíamos, los principios de la escuela fenetista no son los realmente aceptados como criterio de selección de rasgos para las clasificaciones biológicas cuyo objetivo sea reconstruir el árbol de la vida, la metodología fenetista ha tenido un impacto innegable en la taxonomía biológica. Además, dichos métodos son indispensables todavía para casos en los que se desconoce por completo la genealogía de las especies a comparar o en las subdisciplinas en las que se pone un especial énfasis en la caracterización sincrónica de las especies (e.g. morfología). Estos aspectos son los que introducen dificultades mayores a la hora de evaluar qué ocurre con la similitud en biología. Algo análogo ocurre con los modelos psicológicos de las ciencias del color. En los modelos psicológicos, los colores suelen representarse como puntos en un espacio con varias dimensiones (tono, saturación, brillo) y la similitud entre los colores se mide por su proximidad en dicho espacio. La concepción de la similitud que opera en las ciencias del color es también la geométrica (Niiniluoto, 1987).

Así pues, en las disciplinas científicas en las que la similitud genérica se utiliza como un concepto teórico importante (y dejando a un lado el enrevesado caso de la biología), éste no ha desaparecido sino que se trata de la similitud clásica bajo su formulación geométrica como proximidad en algún espacio. Así que creo que la afirmación de Quine(*) de que la similitud genérica está siendo reemplazada progresivamente por otras nociones más específicas de los paradigmas teóricos de las diferentes disciplinas científicas es exagerada¹³. Además cabe argumentar que incluso aunque la noción de similitud genérica estuviera siendo reemplazada progresivamente por otras más específicas eso no haría imposible la elaboración de una teoría general de la similitud que abarcara todas esas nociones como casos particulares.

13. En (Decock; Douven, 2011) se presentan y comparan la teoría de los espacios conceptuales de P. Gärdenfors (que utiliza un modelo geométrico de la similitud) y el modelo de contraste de Tversky para mostrar que la supuesta acientificidad de la noción de similitud no es tal.

2.1.2. Críticas de Goodman

En (Goodman, 1953) y (Goodman, 1972) se argumenta que la similitud es una noción incapaz de cumplir algunos de los roles teóricos que los filósofos le han asignado. En particular, la concepción clásica de la similitud está errada al establecer la relación señalada entre la similitud y las propiedades de los objetos:



“Seven Stricture: Similarity cannot be equated with, or measured in terms of, possession of common characteristics. (...) When in general are two things similar? The first response is likely to be: ‘When they have at least one property in common’. But since every two things have some property in common, this will make similarity a universal and hence useless relation. (...) Are two things similar, then, only if they have all their properties in common? This will not work either, for of course no two things have all their properties in common. Similarity so interpreted would be an empty and hence useless relation. (...) More to the point would be counting not all shared properties but rather only important properties-or better, considering not the count but the overall importance of the shared properties. (...) But importance is a highly volatile matter, varying with every shift of context and interest, and quite incapable of supporting the fixed distinctions that philosophers so often seek to rest upon it. (...) But when to the statement that two things are similar we add a specification of the property they have in common, we again remove an ambiguity; but rather than supplementing our initial statement, we render it superfluous. For, as we have already seen, to say that two things are similar in having a specified property in common is to say nothing more than that they have that property in common. Similarity is not definitionally eliminated here; we have neither a definiens serving as an appropriate replacement for every occurrence of ‘is similar to’ nor a definitional schema that will provide an appropriate replacement for each occurrence. Rather we must search for an appropriate replacement in each case; and ‘is similar to’ functions as little more than a blank to be filled.” (Goodman, 1972)

Así que podemos extraer las siguientes críticas de la cita de Goodman:

- (1) Si definimos la similitud como la coincidencia en alguna propiedad, puesto que todo par de objetos tienen una propiedad en común, todo par de objetos son similares entre sí, con lo cual la similitud es una relación trivial.
- (2) Si definimos la similitud como la coincidencia en todas las propiedades, puesto que para cada par de objetos hay alguna propiedad que uno tiene y el otro no, no habrá ningún par de objetos similares entre sí, con lo cual la similitud es una relación vacua.

PRIMER PREMIO: Simetría y similitud: la concepción transformacional de la similitud



(3) Si decimos que dos objetos son similares en el respecto P o porque tienen la propiedad P , entonces no habremos *definido* la similitud. En cada caso “ x e y son similares” significará algo diferente (dependiendo de la propiedad P que estemos considerando) porque no hay una regla general que nos diga por qué respecto P sustituir la similitud en cada caso, con lo cual la noción de similitud está vacía de contenido.

(4) Si definimos la similitud como la posesión de propiedades importantes, destacadas o relevantes, entonces la similitud será una noción relativa a contextos e intereses.

La objeción (2) no daña a la concepción clásica. Dos objetos similares no tienen por qué coincidir en todas sus propiedades (i.e. no tienen por qué ser indiscernibles). Si definiéramos la relación de similitud como la coincidencia en todas las propiedades esto haría de la similitud una relación necesariamente transitiva y por tanto una equivalencia. Pero generalmente la similitud no es transitiva, como muestran los típicos ejemplos de gradaciones de tonos de colores. La indiscernibilidad es sólo un tipo muy extremo de similitud.

(3) supondría definir “ x e y son similares” como “ x e y tienen la propiedad P ” o “ x e y son P ”. Ahora, si “ x e y tienen la propiedad P ” significa “ x e y tienen alguna propiedad P ”, donde la variable ‘ P ’ está ligada por el cuantificador existencial, entonces (3) es (1). En ese caso estamos diciendo que dos objetos son similares si tienen al menos alguna propiedad en común (pero podrían tener varias). Sin embargo, entiendo que no es esto lo que Goodman quiere decir. Parece que dice más bien que “ x e y tienen la propiedad P ” significa “ x e y tienen exactamente la propiedad P ”, es decir, x e y son similares si tienen esa única propiedad P (al menos en ese contexto). Pero esta relación es una equivalencia, y por ende es necesariamente transitiva. La similitud no es la equivalencia en el respecto P , o la coincidencia en la propiedad P , sino la equivalencia en *algún* respecto P o la coincidencia en *alguna* propiedad P . Eso es lo que permite que dos objetos puedan tener varias propiedades en común o ser equivalentes en varios aspectos diferentes. Y eso es lo que hace que la similitud no sea necesariamente transitiva, dado que x e y podrían ser equivalentes en un respecto P , y e z ser equivalentes en un respecto Q , pero x e z no ser equivalentes en ningún respecto.

Respecto a (1), puesto que Goodman asume que decir que dos objetos tienen una propiedad en común es equivalente a decir que caen bajo un mismo predicado, esto parece comprometerle a una concepción *abundante* de las propiedades por la cual (casi) cualquier predicado designa una propiedad (Niiniluoto, 1987). Esto hace fácil manufacturar propiedades compartidas por cualquier par de objetos, como por ejemplo propiedades “disyuntivas” o propiedades “negativas”. Por ejemplo, una oveja y un cajón de madera serían similares por satisfacer el predicado “no ser un árbol”, que expresaría una propiedad negativa. También serían

similares por satisfacer el predicado “ser una oveja o un cajón de madera”, que expresaría una propiedad disyuntiva. Este supuesto también conduce al “Ugly Duckling Theorem”, según el cual la cantidad de predicados satisfechos por dos objetos cualesquiera es constante. Es decir, un par de objetos cualesquiera tendría exactamente el mismo número de propiedades en común que otro par cualquiera, con lo cual tampoco se podría definir una similitud tetrádica $Q(x,y,u,w)$ del tipo “ x e y son más similares entre sí que u e w si x e y tienen más propiedades en común que u e w ”. En este caso también hay que suponer que el conjunto total de propiedades (i.e. predicados) considerado está cerrado bajo intersección, disyunción y negación (Niiniluoto, 1987).

Pero ésta no es la única opción, no hay por qué aceptar que la negación o disyunción de propiedades sea una propiedad. Podríamos asumir una concepción escasa [sparse] de las propiedades por la cual sólo algunos predicados corresponden a propiedades reales, por ejemplo, como defiende (Armstrong, 1978)¹⁴. Esto es, podríamos sostener, por ejemplo, que una propiedad es relevante si hay alguna teoría científica actualmente aceptada que postule su existencia. O podríamos reforzar la condición exigiendo que la propiedad fuera indispensable para dar cierto tipo de explicaciones teóricas (e.g. causales). Pero incluso aunque tomemos esta opción el problema puede persistir. Dos objetos pueden tener varias propiedades naturales en común y que sólo algunas de ellas sean relevantes al compararlas. Todo esto nos fuerza a enfrentarnos a (4), al fin y al cabo, seleccionar las propiedades ‘naturales’ es una forma de distinguir entre las propiedades importantes y las que no lo son.

Podemos llamar a (4) *el problema del contexto* porque indica que, ante la ausencia de un criterio que nos permita señalar algunas propiedades como más importantes o fundamentales que otras sobre las que basar nuestros juicios de similitud, la selección de las propiedades relevantes depende del contexto y de los intereses de aquel que emite el juicio. Creo que (4) es un problema de difícil solución. Claramente, el contexto contribuye a la selección de ciertas propiedades, la cuestión es si es posible decir algo más al respecto. Dado que en algunos casos sí sabemos qué propiedades son las que hacen que consideremos a dos objetos como similares, podemos intentar dar con ciertos principios o reglas generales que nos indiquen qué propiedades tenderíamos a seleccionar en cada situación, y que expliquen por qué seleccionamos esas y no otras.

14. Como dice (Niiniluoto, 1987), los nominalistas también pueden excluir las ‘propiedades’ disyuntivas y negativas para escapar del Ugly Duckling Theorem. Aunque él habla del nominalismo de predicados, un nominalista de similitud también puede hacerlo. Éste definirá las propiedades como clases de objetos similares entre sí que satisfacen ciertas condiciones (e.g. ser círculos de similitud). Las propiedades negativas y disyuntivas serían otras clases obtenidas a partir de éstas que incumplirían esas condiciones, por ejemplo, sus miembros no tendrían por qué ser todos ellos similares entre sí. Así que no es estrictamente necesario adoptar una posición realista para responder a Goodman.

PRIMER PREMIO: Simetría y similitud: la concepción transformacional de la similitud



El problema del contexto se presenta igualmente en otras disciplinas, así que conviene examinar qué estrategias se han tomado en éstas para abordarlo.

En la biología, por ejemplo, la selección de las propiedades *teóricamente* relevantes supone un problema para la posición taxonómica fenetista (Ereshefsky, 2004). Los taxónomos fenetistas consideran que la exclusión de ciertos rasgos al realizar la comparación por similitud entre las especies es un sesgo injustificado, porque la clasificación ha de ser 'teóricamente neutral'. Esto es porque las clasificaciones han de ser objetivas, repetibles, permanentes y aplicables por cualquier subrama de la biología. La clasificación final sería aquella a la que irían convergiendo las sucesivas clasificaciones que fueran añadiendo un número mayor de rasgos. No obstante, ya hemos señalado que las posiciones dominantes son la evolucionista y la cladista, y para éstas la selección de las propiedades es clara. Sólo las homologías cuentan como propiedades relevantes, porque son las propiedades que proporcionan indicios para descubrir las relaciones genealógicas entre las especies, que es el objetivo de la teoría de la evolución. Así que aquí las propiedades relevantes son aquellas que permiten clasificar a los objetos de acuerdo con las tesis nucleares de la teoría principal.



www.solofici.org

En la psicología cognitiva también se han topado con el problema del contexto. En su caso, han buscado varios principios generales que muestren cómo el contexto contribuye de una manera sistemática a seleccionar las propiedades (psicológicamente) relevantes de los objetos sobre los que se emiten los juicios de similitud. Las propiedades destacadas pueden explicarse apelando al grado de 'prototipicalidad' de los objetos que las poseen (Tversky, 1977), a determinados cambios en el dominio de objetos estudiados (Tversky, 1977), a la mayor/menor intensidad de las propiedades en el contexto (Tversky, 1977), a su capacidad de permitir al sujeto realizar clasificaciones eficientes del dominio de objetos (Tversky, 1977), a su capacidad para permitir que el sujeto que emite los juicios de similitud realice inferencias analógicas informativas y transfiera dicho conocimiento de unos objetos a otros (Gentner, 1983), a su relevancia para formar parte de leyes o principios generales que permitan dar explicaciones del dominio de objetos (Gentner, 1983), a si el sujeto se ha fijado en propiedades del mismo tipo al realizar juicios de similitud sobre dichos objetos (Medin, Goldstone, Gentner, 1993), etc. Esto es, la siguiente batería de principios ha sido propuesta por los psicólogos para acotar los efectos del contexto en la selección de propiedades relevantes:

Principio de Diagnosticidad (Tversky, 1977): dado un dominio de objetos, las propiedades de los objetos que destacan son aquellas que permiten realizar clasificaciones eficientes (e.g. no triviales) del dominio de objetos.

Efecto de Extensión (Tversky, 1977): si todos los objetos del dominio comparten la misma propiedad, ésta no destaca porque no permite clasificarlos eficientemente. Si cambiamos un objeto del dominio por otro que no tenga esa propiedad, entonces esa propiedad destacará porque permite clasificarlos eficientemente.

Efecto de Intensidad (Tversky, 1977): las propiedades que tienen una intensidad muy alta o más alta que el resto (e.g. luminosidad, brillo, volumen) destacan más.

Principio de Inferencialidad/Analogía (Medin et al, 1993): las propiedades de los objetos que nos permiten realizar inferencias analógicas informativas para transferir nuestro conocimiento sobre un objeto conocido a otro desconocido destacan más.

Principio de Sistemática (Gentner, 1983): las relaciones relevantes son aquellas que forman parte de leyes o principios más generales que explican el comportamiento de los objetos del dominio.

Efecto de Prototipicalidad (Rosch, 1988): las propiedades de los objetos que son más prototípicos destacan más.


Efecto de no-independencia (Goldstone; Medin; Gentner, 1991) si el sujeto se ha fijado en la coincidencia/diferencia en propiedades (o relaciones) entonces las coincidencias/diferencias del mismo tipo, i.e. en propiedades (o relaciones), destacan más que las coincidencias/diferencias del otro tipo, i.e. en relaciones (o propiedades).

Estos principios, que poseen apoyo experimental, muestran que si bien seleccionamos las propiedades relevantes al emitir juicios de similitud, dicha selección no es completamente arbitraria sino que está regida por algunos principios generales (Medin et al, 1993). Es interesante observar que la mayoría de estos principios indican que las propiedades relevantes son aquellas que permiten *teorizar* acerca del dominio de objetos, ya sea porque permiten clasificarlos (diagnosticidad), compararlos (intensidad), realizar inferencias analógicas (analogía) o dar explicaciones por medio de leyes o principios generales (sistemática). Es decir, sólo dos de ellas parecen sesgos psicológicos difícilmente generalizables a otras situaciones (prototipicalidad y no-independencia), con lo cual algunos de estos principios podrían plantearse como formas de

PRIMER PREMIO: Simetría y similitud: la concepción transformacional de la similitud

responder a la objeción 4) desde un punto de vista filosófico general: las propiedades relevantes son aquellas que permiten teorizar acerca del dominio de objetos que las poseen. Tomemos, por ejemplo, el efecto de extensión. Si todos los objetos que estamos comparando tienen una misma propiedad, ésta no destaca porque no nos permite clasificar a los objetos distinguiéndolos por tipos. Pero si añadimos, quitamos o sustituimos alguno de los objetos por otro que carezca de esa propiedad, ésta pasará a ser relevante, porque ahora sí permite clasificarlos. Ésta es una observación muy simple, pero explica al menos algunos aspectos relativos a cómo escogemos las propiedades relevantes.

En las secciones posteriores presentaré la concepción transformacional de la similitud. Esta propuesta ofrece una nueva forma de seleccionar las propiedades relevantes en la comparación por similitud entre objetos: las propiedades relevantes son las propiedades invariantes bajo ciertas transformaciones asumidas como aceptables o permisibles entre los objetos. Creo que esta condición es particularmente útil para los casos en los que desconocemos las propiedades que hacen que los objetos sean similares entre sí.

(Medin et al, 1993) también señalan que incluso aunque se caracterice la similitud como la coincidencia en alguna propiedad, ese es sólo el primer paso para una explicación completa de la formación de dichos juicios de similitud. Los objetos suelen asemejarse por coincidir en varias propiedades, y una buena teoría de la similitud debería explicar cómo se seleccionan las propiedades, cómo se las compara y cómo se las combina para dar finalmente un juicio de similitud genérico. Para los psicólogos es particularmente importante dar con los procesos cognitivos que están involucrados en esa selección, comparación y combinación de propiedades. E incluso cuando no ahondan en estas cuestiones, como es el caso del modelo geométrico o el modelo de contraste, la explicación de la similitud entre los objetos va más allá de la afirmación de que comparten alguna propiedad y conlleva una explicación de cómo se combinan las propiedades comparadas para resultar en una relación de similitud. 

Hay una última consecuencia que quiero extraer de la discusión sobre las objeciones filosóficas a la similitud. Replicando a las objeciones de Quine he señalado que disponemos ya de una teoría matemática de la similitud. Esto supone que incluso aunque la selección de las propiedades relevantes para la comparación por similitud entre dos objetos dependa del contexto, por (1) las propiedades lógicas de la similitud (y la relación (2) con las propiedades, sean éstas las que sean) no son relativas al contexto. Esto es, una vez se ha dado con las propiedades que definen una relación de similitud (o bien si se ha comenzado directamente con una similitud primitiva), la teoría de las estructuras de similitud nos proporciona una serie de resultados sobre la similitud entre los objetos que no depende ni del contexto ni de las propiedades escogidas. De hecho, lo mismo se aplica a cualquier otro modelo (e.g. los psicológicos) de la similitud,


puesto que una vez se ha escogido ya cuáles son las propiedades a considerar todo lo que de ahí se sigue es independiente de los aspectos a los que Goodman atribuye un sesgo de arbitrariedad.



2.2. Críticas desde la psicología cognitiva 2.2.1. Tversky y el modelo de contraste

Desde la psicología cognitiva también se han lanzado objeciones a la concepción clásica de la similitud. Como se ha dicho, el modelo popular que representa la similitud en psicología es (o era, más bien) el geométrico. Sin embargo, en (Tversky, 1977) se presentan resultados experimentales que violan los tres axiomas de un espacio métrico. El caso que nos interesará aquí es el de los juicios de similitud no-simétricos. El problema principal reside en que por diferentes razones algunos objetos (o más bien, sus propiedades) destacan más que otros a la hora de emitir un juicio del tipo “x se parece a y”. Es decir, en muchos casos en los que el sujeto es conducido a realizar juicios de similitud direccionales, uno de los dos objetos destaca más o le llama más la atención que el otro, de modo que el sujeto tiende a realizar afirmaciones como “x se parece más a y de lo que y se parece a x”. Tversky examina varios fenómenos del estilo, los más interesantes se deben al grado de prototipicidad de los objetos y a los símiles y metáforas.



En psicología cognitiva está bastante asumido desde (Rosch, 1988) que algunos objetos exhiben un mayor grado de prototipicidad que otros. Esto es, algunos objetos tienen un mayor número de propiedades que suelen estar típicamente asociadas al tipo de objeto al que pertenecen, y son por tanto más representativos de esa clase que otros. Así, un colibrí es un pájaro más prototípico que un pingüino. Según Tversky dados dos objetos x e y, si y es más prototípico que x, el sujeto tenderá a centrar su atención en y como ‘referente de comparación’ y emitirá un juicio del tipo “x se parece más a y de lo que y se parece a x”. Una persona tenderá a decir que un pingüino es más similar a un colibrí de lo que es el colibrí al pingüino. Los símiles y metáforas también presentan casos de direccionalidad. De hecho, algunos símiles y metáforas cambian su significado si se invierte el orden de los objetos comparados. El ejemplo típico es el de la metáfora “Ese carnicero es un cirujano”, que adquiere un significado diferente si la invertimos en “Ese cirujano es un carnicero”. Símiles como por ejemplo “Sus ojos son como perlas” producen efectos parecidos. En estos casos estamos introduciendo una direccionalidad a propósito: queremos decir que sus ojos son como perlas, de modo que el interlocutor comprenda que estamos atribuyendo a sus ojos algunas propiedades típicamente asociadas con las perlas (e.g. brillo, pureza, etc.), y no viceversa. Por tanto nuevamente parece haber una violación de la simetría aquí, dado que sus ojos se parecen más a perlas que ciertas perlas a sus ojos. 

PRIMER PREMIO: Simetría y similitud: la concepción transformacional de la similitud

Puesto que la concepción geométrica asume los axiomas de un espacio métrico y por ende la simetría de la similitud, tiene dificultades para representar los juicios no-simétricos. Por ello Tversky introdujo su modelo de contraste, que es capaz de explicar la violación de la simetría en esos casos en los que el sujeto parece dirigir su atención más a un objeto que al otro. El modelo consiste en 5 axiomas para una estructura y un teorema de representación para la misma. Los dos últimos axiomas son instrumentales y el tercero no añade mucho a nuestra discusión, así que presentaremos sólo los dos primeros. Se asume una estructura (S, Q) donde S es el dominio de objetos a comparar, Q es un conjunto de propiedades/ atributos/ rasgos/ características de los objetos de S :

(1) **Matching:** $S(a,b)=F(A \cap B, A-B, B-A)$

(2) **Monotonicidad:** $S(a,b) \geq S(a,c)$ sii se cumple alguna de las siguientes condiciones:

- (i) $A \cap B \supset A \cap C$
- (ii) $A - B \subset A - C$
- (iii) $B - A \subset C - A$



Es decir, por (1) la similitud entre dos objetos es una función de los atributos que tienen en común y aquellos en los que se diferencia cada uno del otro. Por (2) la similitud entre a y b es mayor a la similitud entre a y c sii coinciden en las mismas propiedades y más o se diferencian en menos propiedades (de entre las mismas). Es decir, la similitud entre los objetos crece cuantas más propiedades comparten y menos diferencias tienen.

Tversky prueba un teorema de representación para (S, Q) que permite asignar ciertos valores (números reales) a los conjuntos de rasgos. De ahí se obtiene la siguiente expresión general de la similitud entre dos objetos, que pondera la semejanza entre los objetos como un contraste de propiedades comunes y diferencias:

$$S(a,b) = \theta f(A \cap B) - \alpha f(A - B) - \beta f(B - A)$$

Aquí f es la función que representa los conjuntos de rasgos en \mathbf{R} . Los parámetros θ, α, β sirven para dar más peso, respectivamente, a la coincidencia y diferencias relativas de los rasgos de a y b . Si el sujeto presta más atención a las similitudes entonces la coincidencia de atributos pesará más, si presta más atención a las diferencias entonces la diferencia relativa de atributos pesará más. De hecho es posible que preste más atención a las diferencias de un objeto respecto al otro que viceversa, como ocurre en los casos de violación de simetría. Tversky puede explicar la violación de simetrías mediante la atención prestada por el sujeto:

Hipótesis de la atención: La dirección de la 'asimetría' está determinada por cuál de los estímulos (objetos) destaca sobre el otro, de modo que el estímulo que menos destaca se parece más al estímulo que más destaca que viceversa.

Esto se puede observar en los parámetros α, β, θ de la expresión general de la escala de similitud, que son los que representan el grado de atención o relevancia de los objetos. Supongamos que dados dos objetos a y b , el objeto b es más prototípico que a , o destaca más. Entonces si preguntamos al sujeto "¿quién se parece a quién?" le obligamos a realizar un juicio de similitud direccional, y lo más probable es que diga que "a se parece más a b (que viceversa)". En el modelo de contraste se puede demostrar lo siguiente (Tversky, 1977):

Proposición (No-simetría): Si $\alpha > \beta$, entonces $S(a,b) > S(b,a)$ sii $f(B-A) > f(A-B)$.

Esto es, el modelo de contraste captura esta diferencia $S(a,b) > S(b,a)$. Si el sujeto se centra en la dirección $a \rightarrow b$, entonces decir que a es más similar a b de lo que es b a a es equivalente a decir que la diferencia $b-a$ es mayor a la diferencia $a-b$. Esto es, en esa dirección $a \rightarrow b$, a es más similar a b (que viceversa) sii b tiene más propiedades que a no tiene (que viceversa) sii a tiene más propiedades que b tiene (que viceversa). En ese sentido, a se parece más a b que al revés porque hay menos diferencias en la dirección $a \rightarrow b$ que en la dirección $b \rightarrow a$.

El juicio de similitud es simétrico sii las diferencias de un objeto respecto al otro tienen el mismo peso o valen exactamente lo mismo:

$$S(a,b) = S(b,a) \text{ sii } \alpha = \beta \text{ o } f(A-B) = f(B-A)$$

Es decir, los juicios simétricos son sólo un caso especial en el que las diferencias de un objeto respecto al otro no pesan más que las diferencias del segundo respecto al primero, es decir, son casos en los que ninguno de los dos objetos 'recibe más atención' que el otro por parte del sujeto. Así pues, en el modelo de Tversky la similitud no es necesariamente simétrica. Esto supone una objeción tanto para la concepción clásica de la similitud como para la geométrica. El defensor de estas concepciones tiene la obligación de explicar los juicios de similitud no simétrica en términos de la similitud simétrica.



PRIMER PREMIO: Simetría y similitud: la concepción transformacional de la similitud

2.2.2. Réplicas a las críticas de la psicología cognitiva

Creo que las críticas de la psicología cognitiva a la concepción clásica no son tan fulminantes como pueden parecer a primera vista. En primer lugar hay cierta tensión que surge si rechazamos la simetría, puesto que entonces tenemos que rechazar también la tesis (2). Eso es bastante problemático. En segundo lugar, la concepción geométrica puede acomodar los casos de no-simetría de varios modos diferentes, sin salirse de la conceptualización de la similitud mediante la distancia en un espacio métrico. Por último, el modelo de contraste tiene sus propios problemas que conviene subrayar.

a. ¿Hasta dónde mediante una similitud no simétrica?

Respecto al primer argumento, podemos preguntarnos hasta dónde podríamos llegar sin adoptar la simetría. Supongamos que adoptamos una similitud no-simétrica. Entonces no podemos asumir (2), porque (2) implica la simetría. Por tanto, si queremos una similitud no necesariamente simétrica tendremos que rechazar (2). Pero los casos de metáfora y comparación indican que no puede violarse totalmente. Supongamos que escuchamos un símil como “sus dientes brillan como el fósforo”. En esta comparación se dice que sus dientes se parecen al fósforo, porque brillan. Ahora bien, independientemente de si sabemos o no algo sobre el fósforo, hay algo que sabemos tras escuchar la comparación: que sea lo que sea el fósforo, éste brilla. Así que, al menos en los casos de comparaciones, hay una propiedad común a ambos objetos aunque la comparación sea direccional. Esto es razonable, porque una de las funciones de la similitud es permitir realizar analogías que transfieren lo que sabemos sobre unos objetos a otros que posiblemente desconocemos (Medin et al, 1993). Por tanto, hay que aceptar algo como la mitad izquierda de (2): si un elemento es similar a otro, entonces ambos tienen alguna propiedad en común. Así pues, tendríamos una situación como la siguiente: x se parece a y , x e y tienen una propiedad en común, podemos inferir información acerca de y a partir de lo que sabemos sobre x , pero aun así y no se parece a x . Esta situación es un tanto sospechosa.



Hay otro diagnóstico diferente. Parece posible decir que dos objetos se parecen el uno al otro aunque uno de ellos se parezca más al otro que viceversa. Es decir, “ x e y se parecen” e “ x se parece más a y que y a x ” son relaciones distintas y compatibles. La primera es absoluta, la segunda es gradual. En ese caso la objeción de Tversky afectaría al modelo geométrico, pero no a la concepción clásica.

b. Juicios de similitud no simétricos en modelos geométricos

En segundo lugar, la concepción geométrica dentro de la propia psicología cognitiva sí puede explicar los juicios direccionales y acomodarlos en sus modelos. Hay varias maneras de hacerlo. Puesto que se asume que la direccionalidad se debe a un cambio en el foco de atención del sujeto que emite el juicio, porque el objeto destaca o le llama la atención, se puede introducir en el modelo geométrico un coeficiente que funciona como sesgo de estímulo. Este coeficiente modifica el resultado final de la medida de similitud pero no está incorporado en la propia relación de similitud (i.e. en la distancia). Eso es lo que se hace en (Nosofsky, 1991), donde se muestra también cómo subsumir el modelo de contraste como un caso especial de un modelo geométrico (si se asume la aditividad de rasgos, sin la cual no se puede hacer mucho). Esta modificación del modelo geométrico permite realizar una enorme cantidad de predicciones que encajan con los casos de violación de simetría. Por otro lado, en (Nosofsky, 2011) se propone el modelo del contexto generalizado. En este caso la solución consiste en estrechar o ensanchar las dimensiones (i.e. los rasgos) del espacio para representar el foco de atención puesto en tal dimensión. De este modo se puede dar cuenta también de los casos de violación de simetría sin salirse de la concepción geométrica.

www.solofici.org

c. Problemas para el modelo de contraste

Por último, el modelo de contraste se enfrenta a varios problemas (Hahn; Chater, 1997):

1. Es una consecuencia matemática del modelo de contraste que algunos objetos se parecen más a sí mismos que otros (Nosofsky, 1991). De hecho, son los estímulos destacados (e.g. los objetos más prototípicos) los que se parecen más a sí mismos que otros. Esto es una violación del axioma de minimalidad de los espacios métricos. Según Tversky hay evidencia empírica también a favor de esta tesis. No obstante, parece un poco difícil de comprender en qué consiste el hecho de que algunos objetos se parezcan más a sí mismos que otros.
2. Otro problema consiste en cómo representar en el modelo de contraste propiedades que tengan valores continuos simplemente mediante conjuntos de elementos.

PRIMER PREMIO: Simetría y similitud: la concepción transformacional de la similitud

3. Una última dificultad consiste en que el modelo de contraste ignora las relaciones y por ende, ignora todos los aspectos estructurales de los objetos. Pero los objetos pueden parecerse no sólo por tener rasgos comunes/diferentes, sino por estar en las mismas relaciones (o en relaciones parecidas) con otros objetos. Para explicar estas semejanzas hay que reconceptualizar las relaciones como cierto tipo de conjuntos de rasgos, y ese movimiento trae sus propios problemas.

La primera objeción es, al menos a primera vista, bastante fuerte. La minimalidad no es sólo una condición interesante por su simplicidad, es que parece que negarla sería algo incomprensible. ¿En qué consiste decir que algunos objetos se parecen más a sí mismos que otros? Sin embargo, alguien podría argumentar lo siguiente. Algunos objetos que son más prototípicos que otros son más simples, perfectos o bonitos. Es decir, algunos objetos tienen lo que se conoce como una 'buena forma' o 'buena figura' (Palmer, 1983). Estos objetos tienden a poseer un alto grado de simetría. De hecho se ha sugerido (Palmer, 1983) analizar cómo de 'buena' es una figura o forma en términos de su grupo de simetrías. Este rasgo de algunos objetos no es cognitivamente irrelevante, afecta a los procesos de reconocimiento, descripción, memoria, etc. Así que se podría decir que hay al menos un sentido en el que los objetos pueden ser más parecidos a sí mismos que otros, i.e. si tienen una 'mejor forma'. No obstante, aunque podamos (si es que podemos) dar sentido a una afirmación así, creo que el modelo de contraste no puede explicarlo. Esto es porque el grado de simetría del objeto no depende de que el sujeto ponga o no más atención en él que en el otro. Tampoco está claro que apelar a las diferencias de un objeto respecto a otro sirva de mucho para explicar porqué se parece más a sí mismo. Uno podría suponer que la simetría es una propiedad de los objetos como otras, pero eso es poco satisfactorio sin introducir una diferencia en los grados de simetría.

La segunda objeción parece difícil de esquivar. Supongamos una propiedad como la longitud, que toma valores continuos. Un objeto que mide 1,10 cm y otro que mide 1,05 cm se parecen más entre sí que lo que se parecen a otro que mide 5 cm. No obstante, eso supone que tienen una propiedad en común. Pero, ¿cuál es esa propiedad en el modelo de contraste? ¿"tener una longitud entre 1,00 y 1,20 cm"? ¿Y si introducimos un cuarto objeto que mida 1,08 cm? Entonces este objeto y el primero se parecerán más entre sí que lo que se parecen al segundo. Pero entonces la propiedad ya no será "tener una longitud entre 1,00 y 1,20 cm", tendrá que ser algo como "tener una longitud entre 1,06 y 1,20 cm". Obviamente podemos repetir este razonamiento. El modelo geométrico tiene aquí una clara ventaja, porque puede representar directamente el valor de la propiedad en cuestión (por lo menos si escoge \mathbf{R}^n como dominio). De hecho el modelo geométrico puede capturar estas diferencias porque establece similitudes no sólo entre los objetos (las n-tuplas) sino entre los valores de sus propiedades (las coordenadas) dado que algunos de estos valores están más próximos entre sí

que otros. El modelo de contraste, que sólo apela a elementos y conjuntos de rasgos de esos elementos, tiene dificultades en representar esas propiedades.

La tercera objeción me parece bastante contundente. El modelo de contraste ignora las relaciones entre los objetos y supone que puede explicar todas las semejanzas en términos de rasgos comunes. Supongamos, por ejemplo, que tenemos los números 2, -2 y 7. Los números 2 y -2 se parecen más entre sí, en este contexto¹⁶, de lo que se parecen a 7. Pero, ¿cuál es la propiedad común a ambos? Lo más razonable es decir, simplemente, que el uno es inverso del otro, y que ninguno de los dos está en esa relación con el 7. Y aquí es difícil reconceptualizar esa relación como una propiedad¹⁷. Algo mismo ocurre con objetos un poco más complicados. No comparamos sólo rasgos de los objetos individuales, comparamos situaciones, eventos, procesos, etc., que involucran objetos en relaciones los unos con otros. También comparamos las partes de unos objetos con las partes de otros por su disposición o estructura interna.



Varios experimentos (Markman et al, 1993), (Hahn; Chater; Richardson, 2003), (Medin et al, 1993), (Goldstone et al, 1991), (Gentner, 1983), muestran que los aspectos estructurales de los objetos no pueden ser ignorados porque irrumpen en una gran cantidad de juicios de similitud. Caracterizarlos simplemente como propiedades o rasgos es insuficiente. Desde el punto de vista de la psicología cognitiva se requiere una teoría sobre el proceso de comparación entre los objetos que explique cómo se forma el juicio de similitud a partir de la comparación y ponderación de determinados rasgos de los objetos, cómo afecta el contexto (incluidos los intereses y el conocimiento previo del que dispone el sujeto) a la selección de determinados rasgos y cómo cambian esos juicios en función de los cambios en los objetos a comparar. En la siguiente sección voy a abordar una de las teorías estructurales de la similitud que ha surgido recientemente, la *Teoría de la Distorsión Representacional*.

16. Si introduyéramos varios números naturales más quizás nos decantáramos por decir que 2 se parece más a 7 que a -2, por ser positivo.

17. En cierto sentido se puede hacer, claro, si definimos ciertas clases a partir de esa relación. Definamos $x \approx y$ si $x=y$ o $y=-x$. Esta relación \approx es una equivalencia, e induce por tanto una partición en \mathbf{Z} cuyas clases de equivalencia son del tipo $\alpha=\{x,-x\}$. Así, tendríamos la siguiente propiedad P común a x e $-x$: x es P si x está en α . De este modo podemos invertir la explicación y decir que 2 y -2 se parecen porque tienen la propiedad de estar en una misma clase α que pertenece al conjunto de todas las clases cuyos únicos miembros son un número entero y su inverso, i.e. las determinadas por \approx . (Gentner, 1983) dice, con un ejemplo similar, que el parecido entre 3 y 6 y el parecido entre 4 y 8 (o el parecido entre los respectivos pares de números) no se puede explicar si no es apelando a la relación "es el doble de". No está claro a qué propiedades del 3 o el 4 habría que apelar para explicar eso.

PRIMER PREMIO: Simetría y similitud: la concepción transformacional de la similitud

3. SIMILITUD TRANSFORMACIONAL EN GEOMETRÍA Y PSICOLOGÍA

La concepción transformacional de la similitud es una caracterización de la semejanza entre los objetos que tiene varios puntos de origen. Algunos de ellos son el *Programa de Erlangen* de Klein en la geometría (Klein, 1924), la teoría de la distorsión representacional de la similitud (Hahn et al, 2003) y las teorías de la percepción tanto ecológica (Hagen, 1986) como de orientación gestalt (Palmer, 1983) en psicología¹⁸.

3.1. El Programa de Erlangen

En (Klein, 1924) el matemático F. Klein propuso su Programa de Erlangen como un modo de organizar las diferentes geometrías que estaban emergiendo en el momento. La idea principal consistía en caracterizar una geometría como el estudio de las propiedades de los objetos que resultaban invariantes bajo cierto tipo de transformaciones. Algunas de las teorías que se examinarán a continuación basan sus ideas en el Programa de Erlangen, así que podríamos considerar a éste como la fuente de inspiración de la concepción transformacional de la similitud.

Dado un espacio π , e.g. un conjunto de puntos, una transformación en π es una biyección $f: \pi \rightarrow \pi$. Es decir, una transformación en un espacio es una correspondencia uno-uno que 'convierte' unos puntos del espacio en otros. Puesto que un objeto en el espacio es un conjunto de puntos, las transformaciones convierten unos objetos en otros. La selección de ciertas transformaciones en lugar de otras implica modificar unas propiedades de los objetos y dejar otras intactas.

Ahora bien, hay muchas transformaciones posibles en el espacio. Por tanto, conviene especificar qué tipos de transformaciones son las realmente interesantes. Por ejemplo, la geometría euclídea es el estudio de las propiedades que son invariantes bajo isometrías, que son las transformaciones que preservan las distancias entre los puntos (Modenov; Parkhomenko, 1965):

Definición (isometría). Sea, π el espacio e $d: \pi \times \pi \rightarrow \mathbf{R}^+$ una función de distancia en π . Entonces una biyección $f: \pi \rightarrow \pi$ es una *isometría* en π sii $\forall x, y, d(x, y) = d(f(x), f(y))$.

18. Versiones de la concepción transformacional pueden encontrarse también en la biología, en particular en la morfología, e.g. en la obra de D'Arcy Thompson y sus seguidores. Por otro lado, la descripción grupo-teórica de la estructura de los cristales y de diferentes propiedades moleculares también sugiere, al menos a primera vista, que ahí opera algo como una concepción transformacional de la similitud. Algo análogo puede decirse de los frisos y otros patrones artísticos. Por último, las operaciones de edición como la inserción, sustitución o eliminación de caracteres, como los utilizados en el reconocimiento de patrones, invitan a un análisis parecido. No obstante, extender la discusión al campo de la biología, la química y otras disciplinas sobrepasaría los límites de este trabajo.

Esto es, las isometrías 'convierten' pares de puntos en pares de puntos que están a la misma distancia que el par original (i.e. los mueven). Hay cuatro tipos principales de isometrías: las traslaciones, las rotaciones, las reflexiones y las reflexiones deslizantes. Observamos que las distancias entre los puntos no se modifican al realizar estas transformaciones. En particular, los lados de las figuras (representadas por conjuntos de puntos) no se alteran al desplazarlas, rotarlas o reflejarlas:

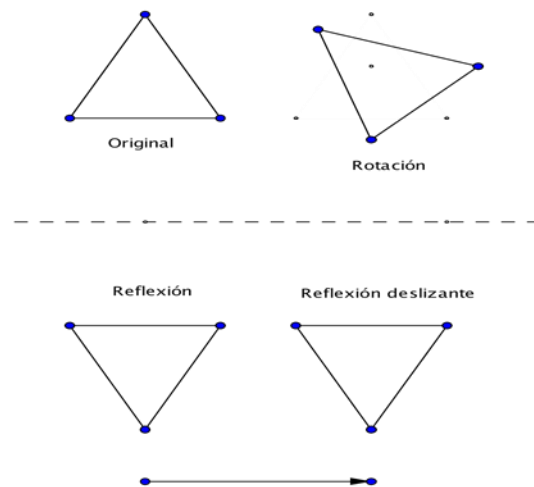


Figura 2. Isometrías sobre el triángulo.

Así, la mayoría de las propiedades intuitivamente asociables con la forma o figura de los objetos quedan preservadas bajo isometrías. Por ejemplo, la longitud de las líneas, los ángulos, el área de las figuras, las proporciones entre segmentos de líneas, etc. son invariantes bajo isometrías. Puesto que la longitud de los lados de las figuras no se modifica, los objetos que pueden obtenerse a partir de otros mediante isometrías tienen el mismo tamaño. Sin embargo la orientación de los objetos no es una propiedad invariante, porque las reflexiones son isometrías, de modo que si sometemos a un objeto a reflexión (como en la imagen) obtenemos un objeto equivalente a pesar de que esté orientado en la dirección contraria.

Pero ésta no es la única geometría posible. Si tomamos otro tipo de transformaciones, obtenemos una geometría diferente. Por ejemplo, podemos tomar todas aquellas transformaciones que preservan la relación triádica Col (A,B,C) de colinealidad entre los puntos, con lo cual obtenemos la geometría afín. Es decir, la geometría afín es el estudio de las propiedades que quedan invariantes bajo transformación es afines (Modenov; Parkhomenko, 1965):

PRIMER PREMIO: Simetría y similitud: la concepción transformacional de la similitud

Definición (Transformación Afin). Sea π el espacio, e $\alpha: \pi \rightarrow \pi$ una biyección. Entonces α es una *transformación afin* en π sii tres puntos A, B, C son colineales sii sus imágenes $\alpha(A)=A', \alpha(B)=B'$ e $\alpha(C)=C'$ son colineales.

Hay varios tipos de transformaciones afines: shears, compresiones/estiramientos (inclinados), reflexiones inclinadas, rotaciones hiperbólicas y elípticas, proyecciones paralelas. En los siguientes ejemplos se puede observar directamente cómo la longitud de las líneas, los ángulos, las áreas y otras propiedades ya no se preservan al transformar a unas figuras en otras bajo transformaciones afines. Se podría decir que si las transformaciones euclídeas corresponden a movimientos rígidos (junto a reflejos), las transformaciones afines corresponden a deformaciones plásticas de los objetos o cambios de perspectiva por parte del sujeto que los observa (junto a reflejos deformados o vistos desde otra perspectiva). Como ejemplos se muestran una reflexión inclinada y una compresión (y su inversa, el estiramiento) en las siguientes figuras.

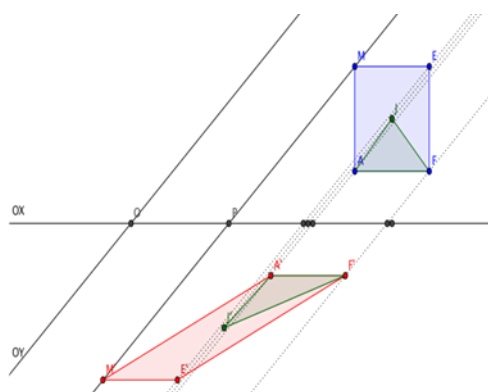


Figura 3. Reflexión inclinada del cuadrado AMEF en el paralelogramo A'M'E'F' y el triángulo AJF en A'J'F' sobre el eje OX en dirección OY. Las distancias y los ángulos no son invariantes afines: la distancia MA es menor a la M'A', y el paralelogramo A'M'E'F' no tiene ya ángulos rectos.



Figura 4. Compresión (rojo) del cuadrado CDMB (para algún $0 < k < 1$) y su inversa (estiramiento verde).

Algunas propiedades de las transformaciones afines son las siguientes (Brannan, 2012):

1. Las imágenes de las líneas rectas son líneas rectas.
2. Las imágenes de las líneas paralelas son paralelas.
3. Se preservan las proporciones de las distancias entre los puntos.
4. Los círculos son equivalentes a/transformables en elipses.
5. Todas las hipérbolas, parábolas y elipses son equivalentes a/transformables en todas las hipérbolas, parábolas y elipses, respectivamente.
6. Todos los triángulos son equivalentes/transformables entre sí.

Esto nos indica algunas propiedades invariantes de la geometría afin, como la 'rectitud' de las líneas, el paralelismo de las líneas, las proporciones de las distancias entre los puntos, los tipos de cónicas (teniendo en cuenta que los círculos y las elipses son afin-equivalentes), etc. Por otro lado, la distancia entre los puntos (y por ende el tamaño de las figuras) o los ángulos no son invariantes afines, como se ha visto. La última propiedad de la lista es muy interesante. Supone que, estrictamente, la 'triangularidad' es una propiedad afin, no 'euclídea'. Esto es porque en la geometría euclídea sí se preservan los ángulos, de modo que triángulos con ángulos diferentes son, *ipso facto*, objetos de tipos distintos. En la geometría afin, sin embargo, se puede probar que todos los triángulos son equivalentes. Así pues, todos los objetos de la figura izquierda son equivalentes entre sí en la geometría afin. Por decirlo de otro modo, la triangularidad no es una propiedad relevante desde el punto de vista euclídeo,

PRIMER PREMIO: Simetría y similitud: la concepción transformacional de la similitud

porque la geometría euclídea permite realizar clasificaciones mucho más finas. Por otro lado la 'cuadrilateralidad' no es una propiedad afín, porque hay cuadriláteros que no son equivalentes entre sí, como muestra la figura derecha. Así pues, la triangularidad y la cuadrilateralidad son propiedades de tipos geométricos diferentes. Así pues una expresión como 'tienen la misma forma' requiere de una mayor cualificación. Debemos precisar que dos objetos tienen la misma forma "en la geometría x".

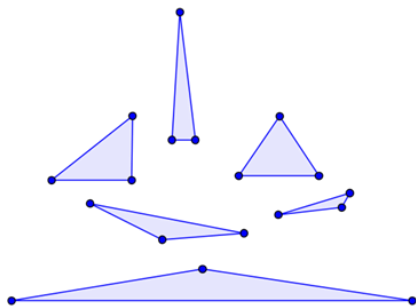


Figura 5. Triángulos afín-equivalentes.

www.solofici.org

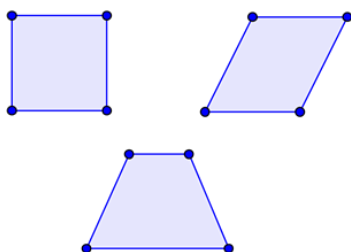


Figura 6. Cuadriláteros no afín-equivalentes.

Se han ofrecido como ejemplos de estudios de diferentes tipos de invariantes el caso de la geometría euclídea y el de la geometría afín. Pero el aspecto realmente interesante de la concepción kleiniana de la geometría reside en la estructura algebraica de dichas transformaciones, que es lo que permite caracterizar completamente cada geometría y clasificarlas. El conjunto de todas las transformaciones sobre el espacio forma un grupo:

Definición (grupo de transformaciones). Sea $(G, *)$ donde G es un conjunto de transformaciones e $*$: $G \times G \rightarrow G$ una operación binaria en G . Entonces $(G, *)$ es un grupo de transformaciones sii:

- (i) Si $f, g \in G$, entonces $f * g \in G$ [Cierre]¹⁹
- (ii) $\exists id \in G$ s.t. $f * id = f = id * f$ [Identidad]
- (iii) $\forall f \exists f^{-1}$ s.t. $f * f^{-1} = id = f^{-1} * f$ [Inverso]
- (iv) $f * (g * h) = (f * g) * h$ [Asociatividad]

Esto es, (i) la concatenación $f * g$ de dos transformaciones, es decir, la realización de una transformación tras otra, es otra transformación de la colección. (ii) Hay una transformación especial, la 'identidad' $id(x) = x$, que deja inalterada cualquier otra transformación con la que se concatena. Es decir, corresponde a la acción de 'dejar las cosas como están'. (iii) Para cada transformación f que realicemos en el espacio hay una transformación inversa f^{-1} , que consiste en deshacer la acción. Por último, (iv) exige que la concatenación sea asociativa. De este modo, cada geometría corresponde a un grupo de transformaciones, que son aquellas que dejan invariantes las propiedades 'propias' de dicha geometría.

Las isometrías que caracterizan a la geometría euclídea forman un grupo de transformaciones, al igual que las transformaciones afines. Para cada traslación, por ejemplo, hay otra traslación en la dirección contraria y la misma distancia, de modo que si movemos un punto bajo la primera traslación y luego realizamos la segunda, volvemos a la posición inicial, i.e. es como si no hubiéramos movido el punto. Del mismo modo, para cada compresión (de un coeficiente k) hay una transformación inversa que es un estiramiento, de modo que aplicar uno tras el otro devuelve a la figura a su apariencia original.

¿En qué sentido es ésta una concepción 'transformacional de la similitud'? La semejanza entre los objetos se caracteriza como la existencia de alguna transformación de unos en otros. Esto es porque cada grupo de transformaciones establece su propio criterio de igualdad o equivalencia entre los objetos del espacio sobre el que actúan:

Definición (G-congruencia). Sea π el espacio y $(G, *)$ el grupo de transformaciones que actúan sobre π . Entonces dos objetos A y B de π son *G-congruentes* sii $\exists f \in G$ s.t. $f(A) = B$.

Debido al elemento de identidad y a la inversa para cada elemento, la G -congruencia es reflexiva y simétrica, respectivamente. Puesto que la composición de dos transformaciones es otra transformación, la G -congruencia es transitiva. Luego la G -congruencia es una relación de equivalencia en π , que organiza así los objetos de π en clases de equivalencia. Dos objetos G -congruentes tienen las mismas propiedades invariantes:

¹⁹ Aunque para que $f * g$ esté definida el dominio de f tiene que ser equivalente al codominio de g , claro. En este caso eso está garantizado porque las transformaciones tienen la forma $f: G \rightarrow G$.

PRIMER PREMIO: Simetría y similitud: la concepción transformacional de la similitud

Definición (Propiedad G-invariante). Sea π el espacio y $(G, *)$ el grupo de transformaciones que actúan sobre π . Sea P una propiedad de un objeto A . Entonces P es una *propiedad G-invariante* sii si B es un objeto G -congruente a A , entonces B tiene la propiedad P .

Así pues, en cierto sentido la equivalencia entre los objetos se debe a las propiedades que comparten. Es decir, dos objetos equivalentes tienen propiedades en común. Sin embargo, la selección de las propiedades *relevantes* la realiza la transformación del grupo en cuestión. Nótese que la igualdad entre los objetos de π es relativa al grupo escogido. Es decir, dos objetos no son simplemente 'iguales', sino G -iguales o iguales en tal o cual aspecto, que está determinado por la transformación del grupo.

De este modo es posible caracterizar una geometría algebraicamente, esto es, por cierto grupo de transformaciones. Dichos grupos de transformaciones están, además, en relaciones los unos con los otros. Esto posibilita a su vez organizar las geometrías mismas, que era el objetivo del *Programa de Erlangen*. Algunos grupos de transformaciones son subgrupos de otros, es decir, están 'incluidos' en ellos, de modo que toda transformación perteneciente al primer grupo es una transformación del segundo grupo y la estructura algebraica se preserva. Por tanto, dos objetos transformables (equivalentes) en la primera geometría siguen siéndolo en la segunda. Puesto que el grupo más general contiene transformaciones que el grupo más pequeño no posee, objetos que no eran transformables (equivalentes) en la primera geometría pueden serlo en la segunda. Por ejemplo, todas las isometrías (traslaciones, reflexiones y rotaciones) son transformaciones afines. Por tanto, el grupo de isometrías de π es un subgrupo del grupo de transformaciones afines de π . Esto supone que todas las invariantes afines son invariantes isométricas, pero la inversa no se cumple. En particular, desde el punto de vista kleiniano la distancia es una propiedad geométrica relativamente poco importante, porque sólo es invariante bajo isometrías. Hay muchas otras propiedades geométricas que no dependen de la distancia.



Es decir, cuanto más general es la geometría, esto es, cuanto más general es el tipo de transformaciones que permite realizar, menos propiedades deja invariantes y las clasificaciones resultantes son cada vez más gruesas. Objetos que bajo una geometría parecían pertenecer a tipos diferentes acaban siendo agrupados en una misma clase bajo una geometría todavía mayor, como es el caso de los círculos y las elipses, que son el mismo tipo de objeto en la geometría afín pero son objetos de tipos diferentes bajo la geometría euclídea.

20. Esto hace de la 'similitud kleiniana' una equivalencia, pero antes hemos dicho que la similitud no es necesariamente transitiva. Podemos decir: dos objetos son similares sii son G -equivalentes en alguna geometría G introducida en el mismo dominio (véase la sección 4.2.).

En la siguiente tabla se presentan varias geometrías diferentes, con ejemplos de sus transformaciones y algunas de las correspondientes propiedades invariantes.



Geometría	Transformaciones	Invariantes
Métrica	Traslación Reflexión (deslizante) Rotación	Distancia Perpendicularidad Área Proporción de longitud Invariantes de similitud
Similitud	Isometrías Homotecia radial	Tamaño de Ángulo Posición del centro de gravedad Invariantes afines
Afín	Similitudes Reflexión inclinada Compresión (inclinada) Rotación hiperbólica Rotación elíptica Shear Proyección paralela	Paralelismo Proporción de división Invariantes proyectivas
Proyectiva	Afinidades Proyección Central	Cross-ratio División harmónica Colinealidad Entre Invariantes topológicas
Topológica	Proyecciones centrales Homeomorfismos	Conexión

Tabla 1. Geometrías e invariantes, tomada (con cambios) de (Hagen, 1986).



PRIMER PREMIO: Simetría y similitud: la concepción transformacional de la similitud

3.2. La Teoría de la Distorsión Representacional (RDT)

En la psicología cognitiva podemos encontrar también teorías que explican la similitud entre objetos en términos de su transformabilidad mutua. Una de ellas es la RDT (Teoría de la Distorsión Representacional) o teoría transformacional de la similitud (Hahn et al, 2003), que se propone como alternativa a los modelos geométrico y de contraste. Un antecedente de la RDT lo encontramos en (Goldman, 1986):

“The contents of representations R and R are similar – receive a high similarity rating- if and only if the content of one is obtainable from the content of the other by application of a higher-ranking cognitive transformation (or combination of transformations)²¹.” (Goldman, 1986)*

En dicho texto Goldman sugiere que los contenidos de dos representaciones mentales son similares si el contenido de una puede obtenerse aplicando algún tipo de transformación sobre el contenido de la otra. Análogamente, la RDT conceptualiza la similitud entre dos objetos como el mínimo número de transformaciones necesarias para convertir a una representación de uno en una representación del otro. Es decir, dos objetos serán más similares cuanto menor sea el número de distorsiones que haya que realizar en una de las representaciones para obtener la otra. La RDT surge como una teoría psicológica de la similitud alternativa a los modelos espacial y de contraste porque, según sus defensores, el problema principal de estos modelos es que no son aplicables a representaciones estructuradas de los objetos, sean éstos oraciones, fragmentos musicales, rostros, voces, animales, etc.

Los defensores de la RDT proponen utilizar la noción de complejidad (condicional) de Kolmogorov, que proviene de la teoría algorítmica de la información, para formalizar la similitud entre representaciones. La complejidad de Kolmogorov de un objeto es algo así como una medida de la cantidad de información necesaria para representar a dicho objeto. La complejidad condicional de Kolmogorov $K(x|y)$ de un objeto x dado otro objeto y y corresponde al programa f más corto necesario para obtener x a partir de y , i.e. para que $f(y)=x$. Mediante la noción de complejidad condicional de Kolmogorov se puede definir una métrica, de modo que la teoría transformacional tiene también una versión geométrica. El objetivo es que la noción informacional sea lo

21. La continuación de la cita también es interesante: *“Notice the analogy between this idea and one employed by topologists. In topology two figures are said to be topologically equivalent (homeomorphic) just in case one can be obtained from the other by some sort of topological operation. If this line of analysis proved promising, it might be exploited by turning it on its head. (I owe this suggestion, on the example that follows, to Charles Chastain.) Instead of studying similarity in terms of transformations, one could study transformations in terms of judged similarities. For example, given perceived similarities among faces, one might try to identify high-ranking transformations by seeing which would account for those similarity judgements.”*

suficientemente general como para medir la similitud entre dos objetos sean cuales sean los medios utilizados para representarlos.

A pesar de que los autores proponen una formalización informacional de la similitud, ésta no realiza el trabajo real en su teoría. El punto fuerte de la propuesta RDT estriba en una serie de experimentos basados en transformaciones sencillas (insertar objetos, invertir sus colores, estirarlos, reflejarlos, etc.) que presentan problemas para el modelo de contraste, que tiene dificultades para realizar, sin muchos supuestos *ad hoc*, las predicciones que sí puede hacer el RDT. Por ejemplo, el primer experimento consiste en presentar a un grupo de sujetos un primer estímulo O , que es una sucesión de círculos llenos o vacíos. Después se le van presentando otras sucesiones de círculos obtenidas a partir de O mediante una o varias de las siguientes transformaciones: insertar/borrar un círculo, desplazar todos los círculos hacia la derecha/izquierda (“cambio de fase”), rellenar/vaciar los círculos vacíos/llenos (“inversión”), reflejar los círculos (“reflexión”).

Se le pide al sujeto que vaya juzgando cuáles de los estímulos posteriores se parecen más al original que otros. Se predice que cuanto menor/mayor sea el número de transformaciones realizadas sobre el original para obtener una determinada sucesión de círculos, mayor/menor será la similitud juzgada por el sujeto entre el original y dicha sucesión de círculos. Los autores muestran que hay una relación estadísticamente significativa entre la similitud percibida y la distancia transformacional entre los objetos, como se esperaba. El modelo de contraste tiene dificultades para hacer estas predicciones, debido a cómo representa los objetos como conjuntos de rasgos. En el modelo de contraste los objetos tendrían que ser las sucesiones de círculos. Respecto a sus rasgos, aquí hay libertad de elección, pero por lo menos deberían reflejar el hecho de que la sucesión de círculos es una lista ordenada de objetos de dos tipos diferentes, unos llenos y otros vacíos. Si elegimos rasgos como “tener tres círculos vacíos” tendremos problemas, porque las sucesiones pueden ordenarse de maneras muy diferentes aun teniendo tres de sus círculos vacíos. Por tanto sus propiedades serían algo como “primer círculo lleno”, “tercer círculo vacío”, etc. Después se podría computar la similitud entre O y cualquiera de sus variantes asignando un valor a los conjuntos consistentes en los rasgos comunes (i.e. los círculos en los que coincidieran por ser vacíos o llenos) y los rasgos diferentes (de cada uno respecto al otro). El problema es que las siguientes dos sucesiones de la izquierda serían máximamente diferentes en el modelo de contraste, pero muy similares en el transformacional. Esto es porque para obtener la segunda a partir de la primera sólo hay que realizar una inversión. Sin embargo, la primera sucesión y la segunda no coinciden en ningún rasgo, porque todo círculo lleno en la primera está vacío en la segunda y viceversa. La teoría transformacional predice, sin embargo, que las dos sucesiones anteriores se parecen más entre sí que las dos de la derecha, a pesar de que éstas sí tienen rasgos comunes:

PRIMER PREMIO: Simetría y similitud: la concepción transformacional de la similitud

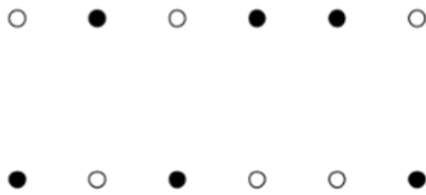


Figura 7. Sucesiones transformables por inversión.

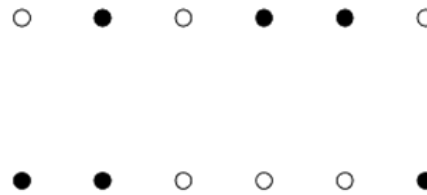


Figura 8. Sucesiones con círculos comunes.

www.solofici.org

El defensor del modelo de contraste podría tratar de buscar otras propiedades de las sucesiones que se pudieran introducir para ajustarse a los resultados experimentales. La cuestión es si puede hacerlo sin que sus movimientos sean demasiado *ad hoc*, y sin apelar indirectamente a su vez a transformaciones en las que estarían basadas esas propiedades. (Hahn et al, 2003) presentan algunos experimentos más en los que muestran las limitaciones del modelo de contraste para dar cuenta de la similitud entre otros tipos de objetos. En muchos de los casos la diferencia en las propiedades de los objetos se identifica sólo una vez se ha detectado la transformación de uno en el otro. Por ejemplo la orientación de un objeto se reconoce como una propiedad destacada del mismo sólo una vez se ha detectado que haría falta rotarlo de una determinada manera para obtener al otro objeto. Así pues parece que al menos desde el punto de vista psicológico, el modelo de contraste es demasiado limitado para explicar los juicios de similitud acerca de objetos mínimamente estructurados. La RDT propone apelar a transformaciones para explicar la similitud entre objetos estructurados, y muestra cómo algunas propiedades destacan una vez se han identificado ciertas transformaciones mediante las cuales obtener a un objeto a partir de otro.

3.3. Teorías transformacionales de la percepción



En (Cassirer, 1944) E. Cassirer trató de establecer un vínculo estrecho entre la teoría de la percepción y la concepción kleiniana de la geometría²². Según Cassirer, el problema epistemológico de la percepción no recibía una solución satisfactoria de la mano de las teorías como la de D. Hume. Si la percepción consistiera simplemente en la sucesión constante de impresiones hay fenómenos que no podríamos explicar, como la constancia de la forma o del color en los objetos.

Este fenómeno perceptual lo experimentamos diariamente. Por ejemplo, aunque observemos un objeto desde distintas perspectivas o bajo diferentes condiciones de iluminación, lo percibimos con la misma forma o color. No percibimos un cambio en la forma de las puertas aunque varíemos nuestra posición, tampoco en el color de las hojas de un libro cuando apagamos la luz. Sin embargo, si la percepción consistiera simplemente en cúmulos de impresiones no podríamos explicar porqué tendemos a asociar ciertos subcúmulos de impresiones con el mismo objeto y otros no, a pesar de que entre los primeros puede haber impresiones lo suficientemente diferentes. Por ejemplo, en el caso del color, aunque una estantería colocada en una pared azul la ensombrece parcialmente, no percibimos la pared como una superficie de dos colores azules, uno más oscuro que otro, ni tampoco la percibimos como seccionada en dos partes. Percibimos la pared como una misma superficie de un solo color, y explicamos la aparente diferencia como un cambio en la iluminación, no como un cambio en las propiedades de la pared.

Según Cassirer parece haber un mecanismo adicional en la percepción que detecta las propiedades de los objetos que son invariantes bajo cierto rango de variaciones o transformaciones asumidas como 'permisibles' por la mente. De este modo percibimos un objeto y lo identificamos como el mismo objeto a pesar de que su forma, tamaño o color (o las correspondientes impresiones) estén siendo sometidos aparentemente a variaciones debido a nuestro cambio de perspectiva o a las condiciones de iluminación. Así que para Cassirer impresiones (o los constituyentes de la percepción que sean) aparentemente diferentes son consideradas por nuestra percepción como equivalentes en cierto grado, en el sentido de ser intercambiables o sustituibles unas por otras sin que se detecte un verdadero cambio en las propiedades del objeto al que representan. Esto es, las impresiones están estructuradas por ciertos principios perceptuales que determinan el rango de impresiones en las que pueden transformarse sin que ello suponga un

22. Algunas de estas ideas están ya contenidas en (Cassirer, 1923), en cierto sentido.

PRIMER PREMIO: Simetría y similitud: la concepción transformacional de la similitud

cambio cualitativo en el objeto al que corresponden. Eso es lo que permite explicar a su vez el cambio cualitativo en los objetos, como transformaciones “no permitidas”. De este modo, la relación entre las nociones de invarianza y transformación proporcionada por la caracterización grupo-teórica de las geometrías ofrece una explicación del fenómeno de la constancia perceptual.

Por ejemplo, la siguiente figura tomada de (Palmer, 1983) es muy simétrica. Esto es porque su grupo de simetría contiene ocho transformaciones: la identidad, tres rotaciones (90° , 180° , 270° en sentido antihorario) y cuatro reflexiones (vertical, horizontal, diagonal izquierda y diagonal derecha), que dejan al punto B fijo²³.

www.solofici.org



La vía abierta por Cassirer ha sido retomada en psicología de la percepción por distintas corrientes. En (Palmer, 1983), de orientación gestalt, se propone un marco teórico que explicaría los fenómenos de constancia de forma, percepción del movimiento, ‘buenas figuras’, agrupación perceptual y efectos de marco de referencia. El concepto central para unificar todos esos fenómenos perceptuales sería su invarianza local bajo el grupo de transformaciones de similitud. Es decir, dichos fenómenos serían invariantes bajo traslaciones, rotaciones, reflexiones y escalamientos (aumento/disminución proporcional del tamaño). Por ejemplo, como veíamos en el caso de Cassirer, la constancia de forma es el fenómeno que consiste en percibir que varios objetos tienen la misma forma o figura a pesar de que haya diferencias en su posición, orientación, sentido o tamaño. O lo que es lo mismo, la constancia de forma es la invarianza de la forma bajo transformaciones que modifican la posición (traslaciones), la orientación (rotaciones), el sentido (reflexiones) o el tamaño (escalamientos). Así, el fenómeno perceptual de la constancia de forma se explica como un proceso cognitivo de detección de invariantes en algunas propiedades (i.e. la forma) de los objetos percibidos.

Otro ejemplo es el caso de las ‘buenas figuras’. Nuestra mente tiende a considerar algunas figuras (e.g. un círculo) como más bonitas, simples, regulares, ordenadas o perfectas que otras (e.g. una sucesión aleatoria de puntos y segmentos de línea). Muchos mecanismos cognitivos como el reconocimiento, el recuerdo, la alineación, la descripción o la generación de figuras (o en general, de patrones) dependen de cómo de buenas sean esas figuras. Una de las explicaciones de lo buena que es una figura apela a sus simetrías, cuanto más simétrica es una figura en comparación con otras, mejor es. Ahora bien, la simetría de una figura se puede caracterizar mediante su grupo de simetría, esto es, el conjunto de todas las permutaciones de los componentes de la figura bajo las cuales la figura queda intacta. O lo que es lo mismo, dada una figura cualquiera, consideremos todas las figuras diferentes a ella que podemos obtener sometiéndola a transformaciones como rotaciones o reflexiones. Entonces, cuantas menos variantes transformacionales tenga una figura, más simétrica será.

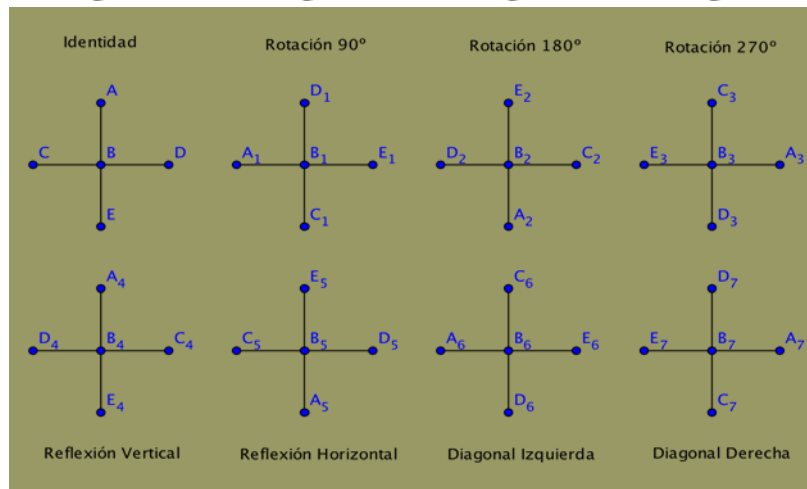


Figura 9. Grupo de simetrías de la cruz.

De este modo, algunas figuras serán mejores figuras que otras cuanto más simétricas sean, esto es, cuanto mayor sea el número de transformaciones bajo el que son invariantes²⁴. Por tanto, el concepto de invarianza bajo las transformaciones de similitud proporciona un marco bajo el cual explicar todos estos fenómenos perceptuales, de un modo análogo a como había propuesto Cassirer.

Por otro lado, (Hagen, 1986) aplica la teoría ecológica de la percepción desarrollada por J.J. Gibson a la teoría de las representaciones artísticas. La teoría ecológica de la percepción sostiene que la percepción del entorno es la localización de invariantes en la luz. En la teoría ecológica percibir es interactuar con el entorno extrayendo información sobre las propiedades invariantes o persistentes (forma, tamaño, distancia, inclinación/pendiente, color) de los objetos sobre cuyas superficies se refleja la luz. Esto es, la luz reflejada queda estructurada de distintos modos en función de las superficies sobre las que se refleja. A medida que el observador se mueve

23. El grupo de simetría de esta cruz es equivalente al grupo de simetrías del cuadrado.

24. Esto es una simplificación. El propio Palmer señala que lo relevante no es el número de simetrías sino cuáles sean esas simetrías y cómo interactúan entre sí.

PRIMER PREMIO: Simetría y similitud: la concepción transformacional de la similitud

alrededor de dichas superficies (y de los objetos que las poseen) va detectando los aspectos de la estructura de la luz que son invariantes bajo los sucesivos cambios de perspectiva o de condiciones de iluminación. De este modo el observador adquiere información sobre las propiedades de los objetos a partir de las propiedades invariantes en la estructura de la luz.

Partiendo de esta teoría de la percepción, el objetivo de Hagen es clasificar y comparar obras pictóricas de diferentes épocas y culturas mediante el tipo de transformaciones geométricas que se han de realizar en las escenas representadas para obtener el contenido de los cuadros que las representan. Estas transformaciones serían simplemente los cambios asumidos por nuestro aparato perceptual como 'permisibles' o no-modificadores de ciertas propiedades presentes en la luz que se refleja en las superficies de la escena. De este modo se hace posible una comparación transcultural de distintos estilos artísticos en un marco psicológico unificado. Según Hagen, la teoría convencionalista de la representación pictórica (de N. Goodman), que sostiene que los cuadros representan las escenas convencionalmente porque son símbolos culturalmente adquiridos por sus usuarios (que han aprendido a utilizarlos e interpretarlos), es poco plausible y además incompatible con la evidencia empírica. La teoría alternativa es la teoría de la similitud, que sostiene que el cuadro representa la escena porque se le asemeja, y dicha semejanza se fundamenta en la percepción visual del sujeto que representa la escena. Pero esta teoría requiere ser refinada: el cuadro representa la escena porque su contenido comparte con ésta ciertas propiedades, que son algunas de las propiedades invariantes de la escena que han sido seleccionadas por el artista que la percibe:

"Invariants theorists argue that pictures succeed as representations of the world because they contain some of the same kind of invariant information in reflected light as the objects of the scene depicted. James Gibson was the mayor advocate of this theory". (Hagen, 1986 p.94)

Así que una representación pictórica es un mapeo de cierto (s) plano(s) (e.g. las superficies de los objetos del mundo) en otro (e.g. el cuadro) que preserva algunas propiedades de los primeros en el segundo. Las diferencias entre mapeos corresponden a transformaciones de geometrías distintas. Tomando otros aspectos como variables, como el número de imágenes que se hacen del objeto (única o múltiples), la distancia del sujeto que pinta respecto a la escena o distintas relaciones entre el plano en el que se representa y la superficie representada, Hagen clasifica todas las obras pictóricas en cuatro estilos artísticos: el europeo (geometría proyectiva), el egipcio (geometría ortogonal), el indio-noroeste (geometría afín) y el japonés (geometría afín).

La noción de similitud que Hagen maneja en todo momento es la transformacional:



"I believe that if people were given a large variety of objects to order in terms of similarity to each other, their order would be determined by the number of geometric invariants shared. The greater the number of shared invariants, the greater the degree of perceived similarity, even when the invariants are too abstract for naive expression. Two objects will be judged most similar to each other when they share the Metric invariants in addition to all the other more abstract ones." (Hagen, p.67)

La representación pictórica y la escena que ésta representa se asemejan porque comparten ciertas propiedades que son invariantes bajo las transformaciones a las que la escena puede ser sometida, que en última instancia se deben a cambios de iluminación o cambios en la perspectiva del observador. Caracterizar la similitud entre la obra representante y la escena representada en términos de transformaciones geométricas permite a Hagen realizar una clasificación lo bastante general como para abarcar obras de culturas muy distintas, pero con el suficiente contenido como para que sea informativa.

www.solofici.org

4. LA SIMILITUD TRANSFORMACIONAL Y LA SIMILITUD CLÁSICA

4.1. Formulación de la concepción transformacional de la similitud



En las secciones previas he presentado diferentes teorías que hacen uso de una caracterización de la similitud en términos de transformaciones. Creo que esos ejemplos muestran que tal caracterización de la similitud tiene un hueco en las disciplinas científicas, en particular en las matemáticas.

Intentemos ahora extraer las tesis principales de dicha "concepción transformacional de la similitud". La concepción transformacional caracteriza la similitud entre los objetos como dependiente de las transformaciones de unos en otros. Cuanto más fácil sea transformar a uno en el otro, sea porque hay un menor número de propiedades que modificar, sea porque dichos cambios son relativamente simples o más específicos, más se parecerán los objetos. Las transformaciones nos permiten identificar fácilmente las propiedades invariantes comunes, es decir, aquellas propiedades que no se modifican al convertir a un objeto en el otro. Son esas propiedades las que consideramos como fundamento de la similitud entre los objetos. Es decir, en lugar de preguntarnos "¿en qué (respecto) se parecen x e y ?", podemos preguntarnos "si convirtiéramos a x en y de esta manera f , ¿qué propiedades de x quedarían sin modificar en y ?"

www.solofici.org

PRIMER PREMIO: Simetría y similitud: la concepción transformacional de la similitud



Ahora bien, ¿sobre qué actúan las transformaciones? ¿Sobre los individuos o sobre sus propiedades? Supongamos, por ejemplo, que veo la sombra de un gato asomándose por detrás de la puerta de mi habitación. A partir de la forma de la sombra y de sus cambios infero que se trata de mi gato. Mi inferencia está justificada por la similitud que hay entre mi gato y su sombra. En este caso, podemos decir que la similitud entre mi gato y su sombra se puede detectar proyectando la forma tridimensional de mi gato en una superficie (el suelo). La forma que obtenemos es la forma de la sombra, que es bidimensional. No tendría mucho sentido decir que la similitud entre mi gato y su sombra se debe a que podríamos transformar fácilmente (mediante una proyección) a uno en el otro. Más bien, lo que parece que sí tiene sentido es decir que podríamos transformar la forma de uno en la forma del otro. Por tanto, parece que las transformaciones actúan sobre las propiedades como la forma, y sólo derivativamente sobre los individuos que las poseen. Esto es, decir que podríamos transformar a un individuo en otro es sólo una forma compacta de decir que podríamos transformar alguna de sus propiedades en alguna de las propiedades del otro. Por ende, dos individuos son similares si tienen propiedades similares, esto es, transformables. Esto complica un poco las cosas, porque ¿cuáles son ahora las propiedades invariantes? Las propiedades invariantes son *propiedades de propiedades*. En este caso, las invariantes son propiedades de las formas (que a su vez son propiedades de los individuos)²⁵.

En segundo lugar, la transformación ha de ser necesaria para que las dos propiedades en cuestión sean equivalentes. Por ejemplo, en el caso anterior, la transformación que necesitamos para que ambas formas sean equivalentes es la proyección de una en la otra. Puede haber varias transformaciones que hagan que ambas sean equivalentes, por ejemplo, una deformación continua de una forma en la otra también haría que ambas fueran equivalentes.

Así que formulemos la primera tesis:

(T1) El criterio para determinar si dos objetos son similares es mostrar que hay una transformación de alguna de las propiedades de uno de los objetos en alguna de las propiedades del otro. Esto es, hay alguna transformación tal que si se realiza en una de las propiedades del primer objeto, hace que ésta sea equivalente a otra de las propiedades del segundo objeto.

El escenario de la sombra nos permite examinar también en qué consistiría una similitud comparativa transformacional.

25. Otro ejemplo: una maqueta es una buena maqueta de un edificio si al escalar algunas de las propiedades del segundo objeto obtenemos algunas de las propiedades del primero, y viceversa al aplicar la operación contraria. No todas las propiedades del edificio tienen porqué preservarse en la maqueta, una maqueta puede ser buena aunque su color no corresponda al del edificio, o aunque no preserve todos los detalles del mismo.

Supongamos que mi gato duerme estirado y lo ilumino con una linterna perpendicularmente. La forma de la sombra que mi gato proyecta en la pared es muy similar a su forma. Esto es por la dirección de los rayos de luz. Aquí, la transformación que se requiere para pasar de una forma a otra es una "proyección ortogonal". La proyección ortogonal es una transformación isométrica en la que los planos (la pared, el 'plano' en el que se ubica la fuente de luz y los planos que cortan al gato) son paralelos y las líneas que unen los puntos de un plano con los del otro (los rayos de luz de la linterna) son paralelas entre sí también. Como es isométrica, esta transformación deja invariantes muchas más propiedades que otras transformaciones. Por ejemplo, el ángulo de la punta de la oreja izquierda de mi gato es equivalente al de la correspondiente 'oreja' de su sombra. Pero ahora supongamos que ilumino a mi gato con la misma linterna desde la esquina de la habitación en lugar de hacerlo frontalmente. La sombra que se proyecta en la pared es una versión 'alargada' de la forma de mi gato. La transformación que se requiere ahora es una "proyección paralela" que produzca un estiramiento en el objeto, y ésta es una transformación afín. En este caso, a pesar de que los rayos de luz siguen siendo paralelos entre sí, los planos ya no lo son, porque la linterna no alumbraba perpendicularmente a mi gato. Algunas de las propiedades que eran invariantes en el primer caso ya no lo son en el segundo. Por ejemplo, la longitud de cualquier parte del cuerpo de mi gato es considerablemente menor a la correspondiente de la forma de la sombra.

Ahora comparemos ambas situaciones. Consideremos la forma de mi gato, la primera forma de su sombra (proyección ortogonal) y la segunda forma de su sombra (proyección paralela). La primera forma de su sombra se parece más a la forma de mi gato que la segunda. Esto es porque la forma de mi gato comparte más propiedades con la primera sombra que con la segunda. En realidad ocurre algo más fuerte todavía. No sólo tienen más propiedades en común, además la forma de mi gato y la primera sombra comparten todas las propiedades invariantes que la forma de mi gato comparte con la segunda, y otras más (e.g. la longitud de las partes del cuerpo). Esto es porque la proyección ortogonal es un tipo específico de proyección paralela, y preserva por tanto las mismas propiedades invariantes e incluso otras añadidas. La primera forma de la sombra es más fácilmente transformable en la forma de mi gato que la segunda forma de la sombra:

(T3) Un objeto Y se parece más a un objeto X que otro objeto Z si las propiedades de Y son más fácilmente transformables en las propiedades de X que lo que lo son las propiedades de Z²⁶.

26. O más precisamente: (T3) Un objeto Y se parece más a un objeto X que otro objeto Z si la transformación de (una de las propiedades de) Y en (una de las propiedades de) X deja invariantes las mismas propiedades que la transformación de (una de las propiedades de) Z en (esa propiedad de) X, y posiblemente otras más.

PRIMER PREMIO: Simetría y similitud: la concepción transformacional de la similitud

Las transformaciones también permiten clasificar las propiedades sobre las que actúan. Si varias propiedades son transformables unas en otras mediante transformaciones de un mismo tipo, entonces todas esas propiedades tienen, por (2), ciertas propiedades invariantes en común. Podemos decir que todas esas propiedades son de un mismo tipo (e.g. ser euclídea). Así que obtenemos:

(T4) Si las propiedades de varios objetos X_1, \dots, X_n son transformables entre sí (y por ende comparten ciertas propiedades invariantes q_1, \dots, q_m), entonces las propiedades de X_1, \dots, X_n pertenecen a una misma clase o son de un mismo tipo.

Nos falta una última tesis. Pero antes quiero formular una posible objeción.

www.solofici.org

Si la similitud entre los objetos está determinada por las propiedades comunes a propiedades de ambos, ¿para qué íbamos a apelar a las transformaciones? La cuestión es que generalmente no conocemos *todas* las propiedades que hacen que dos objetos sean similares y a veces ni siquiera detectamos las *relevantes*. Otras veces no tenemos muy claro qué propiedades son las que se comparten y queremos descubrirlas (por ejemplo, cuando vemos la cara de alguien y sabemos que nos recuerda a otra persona pero no sabemos por qué²⁷). Esto es bastante típico cuando desconocemos a uno de los dos objetos que estamos comparando, y queremos aun así hacer pequeñas estimaciones. Como dicen (Medin et al, 1993), los juicios de similitud del tipo “x e y se parecen” suelen utilizarse tanto en casos en los que se desconoce alguno de los objetos y se pretende inferir analógicamente la presencia de alguna propiedad (del objeto conocido) en el objeto desconocido, como en casos en los que no se quiere precisar en qué respecto son similares, por ejemplo porque se quiere dejar abierta la posibilidad de que se detecten más propiedades comunes en el futuro (de hecho esta es la característica explotada por las metáforas). Así que a veces hace falta un criterio adicional para descubrir las propiedades comunes.



27. El ejemplo de las caras también lo discute (Armstrong, 1978), porque supone una limitación epistemológica evidente para toda caracterización de la similitud en términos de propiedades. Para Armstrong no hay problema para la definición de la similitud como coincidencia en universales porque que no podamos detectar qué universales comparten las dos caras no implica que no compartan algunos y que éstos no sean la razón por la cual las consideramos similares. Independientemente de que uno acepte la posición de Armstrong, hace falta un *criterio epistemológico* para dar con esas propiedades, yo propongo el transformacional.

Por ejemplo, tomemos el siguiente conjunto de propiedades, que son formas de distintos objetos:

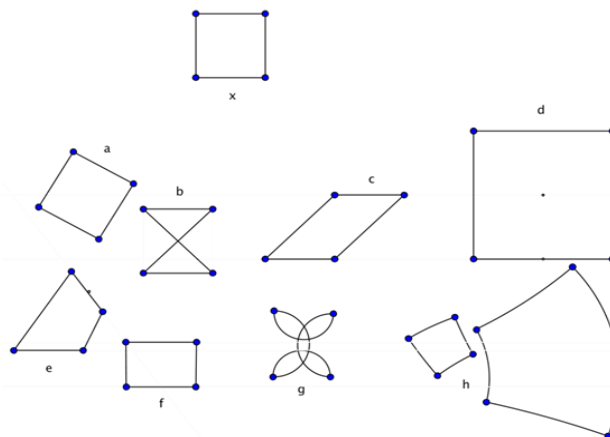


Figura 10. Similitud transformacional entre formas.

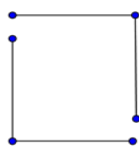
Todas estas figuras se parecen en alguna medida unas a otras. No obstante, algunas figuras se parecen más al cuadrado original, situado en la parte superior, que otras. Aquellas figuras que son el resultado de una deformación menor o ‘más simple’ son más parecidas al original que las demás. De hecho, parece fácil establecer una gradación de similitud entre las figuras utilizando ese criterio. Lo interesante es que requiere algo de trabajo descubrir qué propiedades son comunes a los distintos pares de objetos. Todos ellos, podría decirse, son figuras que poseen ‘cuatro lados’ y ‘cuatro vértices’. O por lo menos, si se requiere que los ‘lados’ sean líneas, podemos decir que son figuras construidas a partir de ‘cuatro puntos’ y ‘cuatro líneas’.

Ahora bien, eso no explica por qué algunas figuras se parecen más entre sí que a otras. Obviamente ello se debe a que comparten más propiedades, pero ¿cuáles son esas propiedades? Una forma de descubrirlo es pensar en qué transformaciones sería necesario realizar en unas para obtener las otras. Consideremos las nueve figuras, las nombramos a, b, c, d las de la fila superior, e, f, g, h ²⁸ las de la fila inferior e x el cuadrado original. Si rotamos ligeramente a x obtenemos la figura a . Esta deformación es la más simple de todas, de hecho, podemos atrevernos a decir que no modifica en absoluto la forma de x . Es decir x e a son la misma forma. Si comprimimos verticalmente a x obtenemos a f . Esta deformación es un poco más sutil que la rotación, porque ahora x y f se parecen menos entre sí de lo que se parecen x y a . Comprobamos que x y f comparten menos propiedades que x y a : x e a tienen lados de la misma longitud, mientras que x y f no. Es más, la proporción entre los lados también ha cambiado en consecuencia. Si inclinamos los lados de x hacia la derecha obtenemos c .

28. En la figura h nombra dos ejemplos de un mismo tipo de transformación (inversión) en x . Se puede pensar en cualquiera de los dos para las explicaciones que siguen.

PRIMER PREMIO: Simetría y similitud: la concepción transformacional de la similitud

Esta transformación también ha modificado algunas características de x que a sí posee: los ángulos internos ya no son ángulos rectos, y las longitudes de sus lados también han cambiado. Sin embargo, tanto f como c comparten con x e a una propiedad interesante, y es que sus lados siguen siendo mutuamente paralelos. Eso distingue a estas cuatro figuras de las demás. No obstante puede que, a pesar de que se trata de figuras bastante simples, el paralelismo de los lados no sea ya una propiedad tan evidente. La búsqueda de propiedades comunes se complica cuanto más permisivas son las transformaciones. Si curvamos a x obtenemos h . En este caso, los lados de h ya no son rectas, y sin embargo h se parece más a x que g , por ejemplo. Esto es porque los ángulos de h siguen siendo ángulos rectos. Respecto a x y g , ¿Hay acaso alguna propiedad común a x y g más allá del hecho de que ambas poseen cuatro vértices y cuatro 'lados'? Al fin y al cabo, x y g son bastante diferentes. Sin embargo x y g son conexos, es decir, podemos transitar de uno de sus vértices a cualquier otro a través de sus lados. Esta propiedad se debe a que los vértices están conectados por lados, por supuesto, pero no se debe al número de lados, esto es, no se debe a que las dos figuras tengan cuatro lados.



Esta propiedad se pone de relieve²⁹ cuando comparamos estas figuras con otras como la presentada en la siguiente figura, que es disconexa:

Figura 11. Forma disconexa.

El siguiente cuadro resume la discusión previa, clasificando los objetos por las transformaciones:

	Transformación	Invariantes
$x \rightarrow a$	Isometría (rotación)	Longitud Invariantes de similitud
$x \rightarrow d$	Similitud (escalamiento)	Ángulos Invariantes afines
$x \rightarrow c$ $x \rightarrow f$	Afín (shear & compresión)	Paralelismo de lados Invariantes proyectivas
$x \rightarrow h$	Proyectiva	Colinealidad Invariantes topológicas
$x \rightarrow e$	Inversiva (inversión)	Ángulos Invariantes topológicas
$x \rightarrow b$ $x \rightarrow g$	Topológica (homeomorfismo)	Conectividad

Cuadro 2. Transformaciones de las formas.

29. Esto es lo que predice (Tversky, 1977), es el efecto de extensión. Hemos introducido un objeto que carecía de una propiedad compartida por todos los objetos del dominio, de ahí que dicha propiedad (la conectividad de la figura) ahora sí destaque.

Así pues, la última tesis es:

(T2) La transformación en cuestión deja invariantes algunas propiedades más específicas de las propiedades de esos objetos que pueden ser *descubiertas* una vez se ha realizado la transformación. Estas propiedades específicas son las que fundamentan la similitud entre los dos objetos. Por tanto, dos objetos X e Y son similares si tienen propiedades tales que tienen propiedades invariantes q_i en común.

En resumidas cuentas, propongo que las tesis nucleares de la concepción transformacional son:

(T1) El criterio para determinar si dos objetos son similares es mostrar que hay una transformación de alguna de las propiedades de uno de los objetos en alguna de las propiedades del otro. Esto es, hay alguna transformación tal que si se realiza en una de las propiedades del primer objeto, hace que ésta sea equivalente a otra de las propiedades del segundo objeto.

(T2) La transformación en cuestión deja invariantes algunas propiedades más específicas de las propiedades de esos objetos que pueden ser *descubiertas* una vez se ha realizado la transformación. Estas propiedades específicas son las que fundamentan la similitud entre los dos objetos. Por tanto dos objetos X e Y son similares si tienen propiedades tales que tienen propiedades invariantes q_i en común.

(T3) Un objeto Y se parece más a un objeto X que otro objeto Z si las propiedades de Y son más fácilmente transformables en las propiedades de X que lo que lo son las propiedades de Z .

(T4) Si varios objetos X_1, \dots, X_n son transformables entre sí (y por ende comparten ciertas propiedades invariantes q_1, \dots, q_m), entonces X_1, \dots, X_n pertenecen a una misma clase o son de un mismo tipo.



4.2. Similitud transformacional como extensión de la similitud clásica



En esta sección se ofrece una pequeña caracterización formal de la concepción transformacional de la similitud para poner de relieve su relación con la clásica. La forma más simple de representar la similitud transformacional entre objetos sería decir que dos objetos x e y de un dominio S son similares si hay una función $f: S \rightarrow S$ tal que $f(x) = y$. Creo que esto permitiría cubrir casi cualquier ejemplo concebible.

PRIMER PREMIO: Simetría y similitud: la concepción transformacional de la similitud

Pero no sería una caracterización muy útil. Los objetos matemáticos son entidades gregarias, y las funciones no tienen porqué ser una excepción. Así que sería más interesante tomar sólo las funciones que cumplan ciertas propiedades e introducir alguna *estructura* en la que dichas funciones puedan habitar. Por ello, propongo seguir las líneas generales del proyecto de clasificación kleiniano, y escoger como hábitats los grupos de transformación. El plan es el siguiente: sea S un conjunto de objetos y \mathfrak{G} una colección finita de grupos de transformación $\mathfrak{G} = \{G_1, \dots, G_n\}$ sobre S , i.e. cada G_i es un subgrupo de S_S , el grupo simétrico de S (el conjunto de todas las biyecciones $f: S \rightarrow S$). Entonces definimos una similitud \sim entre los elementos de S como la existencia de alguna transformación (en algún grupo G_i) de un elemento de S en otro. Cada grupo G_i produce una partición en S diferente por la relación de equivalencia $x \approx_i y$ sii hay una transformación de x en y en el grupo G_i . De modo que la similitud entre elementos corresponde a su equivalencia bajo las transformaciones de *algún* grupo.

Después definimos una relación de similitud triádica en S . Para eso definimos la relación $T(x,y,z)$ “ y se parece más a x que a z ” como “para toda clase de equivalencia en la que ocurren x e z hay una clase de equivalencia más fina en la que ocurren x e y ”. De este modo capturamos la idea intuitiva de que un objeto se parece más a otro que un tercero si es más fácilmente transformable en él o si su transformación en él es más específica.

Por último, se muestra que si asignamos a cada elemento de S el conjunto de las clases de equivalencia a las que pertenece (una por cada grupo), i.e. el conjunto de sus órbitas, dicha función es un *quasianálisis débil no-estándar* de S en el conjunto de todas las clases de equivalencia. Se muestra que la relación triádica $T(x,y,z)$ es equivalente entonces a la relación triádica de (Mormann, 2009) que se da entre las *quasipropiedades* asignadas a dichos elementos, y tiene por tanto las mismas propiedades. De este modo se comprueba que la explicación transformacional es compatible con la concepción clásica de la similitud, dado que la similitud satisface en este caso las dos condiciones expuestas en la sección 1, donde las propiedades quedan interpretadas como las clases de equivalencia inducidas por las transformaciones:

Sea un conjunto S de objetos y una colección finita \mathfrak{G} de grupos de transformación $\mathfrak{G} = \{G_1, \dots, G_n\}$ sobre S . Entonces (Armstrong, 1988):

Definición (subgrupo). Sea $(G, *)$ un grupo de transformaciones e $H \subseteq G$. Entonces $(H, *)$ es un subgrupo de $(G, *)$ sii:

- (i) $id \in H$.
- (ii) $f, g \in H \Rightarrow f * g \in H$.
- (iii) $f \in H \Rightarrow f^{-1} \in H$.



Cada G_i es un subgrupo de S_S , el grupo simétrico de S (el conjunto de todas las biyecciones $f: S \rightarrow S$). La relación ‘es un subgrupo de’ es un orden parcial en \mathfrak{G} . Ahora definiremos la similitud entre los objetos así:

Definición (similitud). Sea S e $x, y \in S$. Entonces $x \sim y := \exists G_i \in \mathfrak{G}$ s.t. $\exists f \in G_i$ s.t. $f(x) = y$.

Proposición. (S, \sim) es una estructura de similitud.

- **Prueba:** sea x un elemento de S . Entonces como id está en todo G_i , e $id(x) = x$, $x \sim x$. Sean x, y elementos de S tales que $x \sim y$. Entonces hay un f en G_i , para algún i , tal que $f(x) = y$. Como f^{-1} está en G_i , porque es un grupo, e $f^{-1}(y) = x$, $y \sim x$. Luego \sim es una similitud en S .

La similitud \sim no tiene porqué ser transitiva, como muestra el siguiente contraejemplo:

- **Contraejemplo:** sea $S = \{x, y, z\}$ e $\mathfrak{G} = \{G_1, G_2\}$, $G_1 = \{f, id\}$, $G_2 = \{g, id\}$ donde $f(x) = y$, $f(z) = z$, $f(y) = x$ & $g(x) = x$, $g(y) = z$, $g(z) = y$. Aquí f y g son cada una su respectiva inversa en los grupos. Por tanto $x \sim y$ e $y \sim z$ pero no se cumple $x \sim z$.

Esto es, la similitud no es necesariamente transitiva porque podría haber tres objetos x, y, z tales que x e y fueran transformables en un grupo, y e z lo fueran en otro, y puesto que se trata de grupos diferentes no tendría porqué haber una transformación de x en z .

Ahora, cada G_i induce en S una partición mediante la siguiente relación de equivalencia $x \approx_i y$ (Armstrong, 1988):

$$x \approx_i y := \exists f \in G_i \text{ s.t. } f(x) = y$$

Para cada x , la clase de equivalencia $[x]_i$, a la que pertenece, para algún G_i , es su *i-órbita*. Esto es, $[x]_i$ es el conjunto de todos aquellos objetos en los que puede transformarse a x bajo el grupo G_i ³⁰. Las particiones se comparan típicamente por su ‘finura’:

Definición (más fina que). Sean $S|_{\approx_i}$ e $S|_{\approx_j}$ las particiones en S determinadas por las equivalencias \approx_i e \approx_j . Entonces $S|_{\approx_i}$ es *más fina* que $S|_{\approx_j}$ sii $\approx_i \leq \approx_j$ sii $\approx_i \subseteq \approx_j$ $\forall A \in S|_{\approx_i} \exists B \in S|_{\approx_j}$ s.t. $A \subseteq B$.

Ahora definimos una relación de similitud triádica. La noción que queremos capturar es la siguiente: un objeto b se parece más a otro a que uno c sii b es más fácilmente transformable en a que c o la transformación de b en a es más específica/menos permisiva que la de c en a . Esto nos permitirá realizar comparaciones más finas.

30. “(...) the apprehension of the particular qua “existence” involves apprehension of the possibilities of transformation which it contains within itself.” (Cassirer, 1944)

PRIMER PREMIO: Simetría y similitud: la concepción transformacional de la similitud

Como auxiliares definimos los siguientes conjuntos $S(x) = \{[z]_i \mid x \in [z]_i\}$ y $S_{xz} = \{A \in S(x) \mid z \in A\}$. Es decir, $S(x)$ es el conjunto de todas las órbitas de x . S_{xz} es simplemente el conjunto de las órbitas de x en las que está z . La relación entre ambos es muy simple:

$$S_{xy} = S(x) \cap S(y)$$

- **Prueba:** (i) sea A en S_{xy} . Entonces x y y están en $A = [\alpha]_i$. Por tanto A está en $S(x)$ e A está en $S(y)$. (ii) Ahora supongamos que A está en $S(x)$ e en $S(y)$. Luego $x, y \in [\alpha]_i = A$. Luego $A \in S_{xy}$.

Ahora definiremos una similitud triádica $T(x, y, z)$ "y se parece más a x que z ".

Definición (similitud triádica). $T(x, y, z) := \forall A \in S_{xz} \exists B \in S_{xy}$ s.t. $B \subseteq A$.

Es decir, y se parece más a x que z sii para toda órbita de x en la que esté z , hay una órbita de x en la que está y que está incluida en ella. Esto es, podemos convertir a x en y mediante el mismo tipo de transformaciones con las que convertimos a x en z , e incluso otras más específicas. La relación $T(x, y, z)$ no es más que la relación triádica entre quasipropiedades que se propone en (Mormann, 2009). Para aclarar esto probamos:

Proposición. $T(x, y, z)$ sii $S(x) \cap S(z) \subseteq S(x) \cap S(y)$.

- **Prueba:** (i) Supongamos $T(x, y, z)$ e sea $A \in S(x) \cap S(z) = S_{xz}$. Luego $\exists B \in S_{xy}$ s.t. $B \subseteq A$. Como $x, y \in B$, $x, y \in A$, i.e. $A \in S_{xy}$. Luego $S_{xz} \subseteq S_{xy} = S(x) \cap S(y)$. (ii) Supongamos que $S(x) \cap S(z) \subseteq S(x) \cap S(y)$. Sea $A \in S_{xz} = S(x) \cap S(z)$, luego $A \in S(x) \cap S(y) = S_{xy}$. Como $A \subseteq A$, $T(x, y, z)$.

La caracterización más precisa de la concepción clásica de la similitud la encontramos en la versión representacional del quasianálisis (Mormann, 2009):

Definición (Quasianálisis débil). Una asignación $r: S \rightarrow \wp(Q)$ es un *quasianálisis débil* sii la asignación r satisface los siguientes axiomas:

- (C1) $x \sim y \Rightarrow r(x) \cap r(y) \neq \emptyset$
 (C2) $r(x) \cap r(y) \neq \emptyset \Rightarrow x \sim y$.

Puesto que el quasianálisis es extensional, se supone que $Q \subseteq \wp(S)$. En particular, el quasianálisis r es *estándar* sii $Q \subseteq SC(S)$. La asignación r se define como $r(x) := \{T \in SC(S) \mid x \in T\}$, donde $SC(S)$ es el conjunto de círculos de similitud de S . Ahora, sea $q: (S, \sim) \rightarrow (\wp(Q), \sim^*)$ definida como $q(x) = S(x)$, donde $A \sim^* B$ sii $A \cap B \neq \emptyset$.

Proposición. La función q es un quasianálisis débil no-estándar en S .

- **Prueba:** (i) Supongamos que $x \sim y$, luego x, y están en $[\alpha]_i$ para algún i . Luego $[\alpha]_i$ está en $S(x)$ e $[\alpha]_i$ está en $S(y)$, i.e. $S(x) \cap S(y) \neq \emptyset$. (ii) Supongamos que $[\alpha]_i \in S(x) \cap S(y) \neq \emptyset$. Entonces $x, y \in [\alpha]_i$. Luego $x \sim y$. Como q tiene la forma $q: (S, \sim) \rightarrow (\wp(Q), \sim^*)$ donde $Q = \wp(S)$ y cumple (i)-(ii), es un quasianálisis débil en S . Ahora, q es no-estándar porque $Q \not\subseteq SC(S)$.

Es decir, en nuestro caso $S(x)$ es el conjunto de quasipropiedades de x , i.e. el conjunto de sus órbitas. Esto es, las quasipropiedades de x representan el resultado de realizar diferentes transformaciones en x . El quasianálisis en x lo representa como el conjunto de los posibles efectos que tendría realizar distintas transformaciones en x . Cada una de las quasipropiedades de x es algo como un espacio de objetos en los que sería posible transformar a x . El quasianálisis es como un mapa que indica todas las posibles transformaciones de x .



www.solofici.org

Para ilustrar la explicación, podemos tomar las formas de la figura 15. Nuestro dominio de objetos son las formas $S = \{x, a, b, c, d, e, f, g, h\}$. Sobre éstas actúan los grupos de transformación registrados en la tabla, es decir, el grupo de isometrías, el grupo afín, el proyectivo, etc. Ahora, tenemos comparaciones como $T(x, a, c)$, esto es, a (cuadrado rotado) se parece más a x (original) que c (cuadrado shear), pero no tenemos $T(x, c, a)$. Esto concuerda con nuestras intuiciones, y se debe a que la forma a está en toda órbita en la que se encuentran x y c , pero hay al menos una órbita en la que están x y a pero no está c . Ésa es la órbita de x (y de a) determinada por el grupo de isometrías sobre S , dado que no hay ninguna isometría que haga equivalentes a x (ni a a) y a c . Eso es claro, porque para transformar a x en c hace falta una transformación afín (un shear). Lo mismo se aplica a las otras comparaciones. Respecto al quasianálisis, por ejemplo, la imagen bajo el mismo de x es $f(x) = \{\{x, a\}, \{x, a, d\}, \{x, a, d, c, f\}, \dots, \{x, a, b, c, d, e, f, g, h\}\}$, donde $\{x, a\}$ es la órbita de x determinada por el grupo de isometrías, $\{x, a, d\}$ es la órbita de x determinada por el grupo de similitudes, \dots , $\{x, a, b, c, d, e, f, g, h\}$ es la órbita de x determinada por el grupo de homeomorfismos (transformaciones topológicas). Cada una de las órbitas de x es una 'quasipropiedad' del mismo, que contiene información acerca de los objetos a los que se parece por ser de un mismo tipo. O lo que es lo mismo, nos dice los objetos en los cuales puede transformarse de acuerdo con cierto tipo de transformaciones aceptadas. En consecuencia, cada órbita muestra aquellos objetos con los que dicho objeto comparte ciertas propiedades invariantes *exclusivas* de ese tipo. La primera órbita de x , $\{x, a\}$, contiene los objetos en los que x es transformable por

PRIMER PREMIO: Simetría y similitud: la concepción transformacional de la similitud

isometría, los objetos isométricos, y que son los que comparten con x ciertas propiedades invariantes (e.g. longitud de lados) que x no comparte con otros objetos³¹.

Lo que hace el quasianálisis es presentar los objetos de otro modo. El quasianálisis toma un objeto, e.g. x , y proporciona una representación estructural $f(x)$ de éste (Mormann, 1994, 2009) (en el sentido de (Swoyer, 1991)). Esta representación es algo así como una descripción enriquecida de x , en este caso, lo caracteriza relacionamente mostrando todas sus posibles transformaciones en el resto de objetos. La semejanza entre los objetos, e.g. $x \sim a$, corresponde estructuralmente (en el sentido fuerte de que todas las propiedades estructurales relevantes se preservan en dicha correspondencia) a su coincidencia en alguna órbita $f(x) \cap f(a) \neq \emptyset$, es decir, consiste en ser transformables en los mismos objetos bajo un mismo tipo de transformaciones, i.e. en pertenecer a un mismo tipo de objetos.

www.solofici.org

5. CONCLUSIÓN

El objetivo de este escrito era, en primer lugar, presentar la concepción clásica de la similitud y una serie de objeciones que se le han lanzado desde la filosofía y la psicología, junto a las réplicas que podrían formularse desde este marco. El segundo objetivo era explorar la concepción transformacional de la similitud, extrayendo sus tesis nucleares a partir de teorías de distintas disciplinas científicas (geometría, psicología) y mostrando sus conexiones con la concepción clásica o estándar de la similitud.

En primer lugar, las críticas filosóficas y psicológicas a la concepción clásica de la similitud no son concluyentes. La existencia de las teorías y modelos formales de la similitud y del uso de dicha noción teórica en psicología refutan las objeciones de (Quine, 1970). De entre las objeciones de (Goodman, 1972) la más problemática concierne a la dependencia contextual de la similitud, pero se han señalado algunos principios generales que pueden utilizarse para mostrar que la selección de las propiedades relevantes en un contexto dado no es completamente arbitraria. Las objeciones de (Tversky, 1977) a la simetría de la similitud pueden contestarse mostrando que realmente no afectan a los truismos de la concepción clásica (en particular a la simetría), que los modelos espaciales pueden acomodar las violaciones de la simetría (Nosofsky, 1991), y que el modelo de contraste no está exento de algunos problemas importantes (Hahn; Chater, 1997).

En segundo lugar, la caracterización transformacional de la similitud se fundamenta en la evidencia empírica que proviene de las teorías psicológicas presentadas (RDT, ecológica, etc.) de (Hahn et al, 2003), (Hagen, 1986) y (Palmer, 1983), y en la explicación grupo-teórica de la similitud geométrica del *Programa de Erlangen* (Klein, 1924). Esta concepción satisface las exigencias impuestas por las críticas de las teorías estructurales de la similitud en psicología a los modelos clásicos (geométrico y de contraste), porque es aplicable a objetos estructurados. Se ha argumentado que la concepción transformacional de la similitud implica las tesis de la concepción clásica. Es decir, tanto que (1) la similitud es una relación reflexiva y simétrica, como que (2) dos objetos son similares si tienen alguna propiedad en común. Se ha considerado que era más apropiado optar por (2)* dos objetos son similares si tienen alguna propiedad en común o propiedades similares. De este modo, la similitud entre propiedades se reduce a su equivalencia en algún respecto, esto es, a la existencia de alguna transformación que haga a una de las propiedades equivalente a la otra. Esto implica que las propiedades similares (estrictamente, equivalentes) lo son por compartir otras propiedades que son invariantes bajo las transformaciones señaladas. De este modo se puede ofrecer una respuesta alternativa al problema (4) de Goodman acerca de cuáles son las propiedades relevantes al comparar a dos objetos por su similitud: son las invariantes bajo ciertas transformaciones. Esta relación puede verse en la caracterización formal de la concepción transformacional mediante las estructuras de similitud y el quasianálisis (que representa a los objetos por sus 'variantes transformacionales').

No obstante, el rol de las nociones de transformación e invarianza en la explicación de la similitud entre los objetos está todavía por esclarecer. Se ha puesto el énfasis en propiedades como la forma o figura de los objetos, pero a primera vista no está claro cómo podría extenderse esta concepción al caso del color. Tampoco está claro si, en otras disciplinas distintas a la psicología y la geometría, la similitud transformacional (en el caso de que haya algo así) tendrá las características extraídas en el presente análisis. En suma, todavía quedan muchos aspectos por explorar en torno a la conexión entre la semejanza entre los objetos y las nociones de transformación e invarianza.

³¹. Otro ejemplo: una elipse se parece más a un círculo que una parábola, porque para toda órbita en la que están el círculo y la parábola (topológica o proyectiva) hay una órbita incluida en la que están el círculo y la elipse (la misma), pero no viceversa. Esto es porque hay al menos una órbita en la que están el círculo y la elipse pero no la parábola (la afín).



PRIMER PREMIO: Simetría y similitud: la concepción transformacional de la similitud

REFERENCIAS

- (1) Armstrong, D. M. (1978): *Nominalism & Realism: Universals & Scientific Realism, Volume I*. Cambridge: Cambridge University Press.
- (2) Armstrong, M. A. (1988): *Groups and Symmetry*. New York: Springer-Verlag.
- (3) Borges J. L. (1997): *Ficciones*. Madrid: Alianza Editorial.
- (4) Brannan, D.A.; Esplen M.F.; Gray J.J. (2012): *Geometry*. Cambridge: Cambridge University Press.
- (5) Brockhaus, K., (1963): *Untersuchungen zu Carnaps Logischem Aufbau der Welt*. Dissertation, Universität Münster.
- (6) Carnap, R. (1923): *Die Quasizerlegung – Ein Verfahren zur Ordnung nichthomogener Mengen mit den Mitteln der Beziehungslehre*. Manuscrito no publicado RC-081-04-01, University of Pittsburgh. (*Quasizerlegung*).
- (7) Carnap, R. (1967 [1928]): *The Logical Structure of the World & Pseudoproblems in Philosophy*. Traducido por Rolf A. George. Berkeley y Los Angeles: University of California Press (*Aufbau*).
- (8) Cassirer, E. (1923): *Substance and Function*. W. C. Swabey, & M. C. Swabey, (Trans.) Chicago-London: The Open Court Publishing Company.
- (9) Cassirer (1944): *The Concept of Group and the Theory of Perception*. *Philosophy and Phenomenological Research*, 5, nº1.
- (10) Decock, L.; Douven, I. (2011): *Similarity after Goodman*. *Review of Philosophy and Psychology*, 2, 61-75.
- (11) Diéguez, A. (2012): *La vida bajo escrutinio. Una introducción a la filosofía de la biología*. Biblioteca Buridán.
- (12) Ereshefsky, M. (2004): *The Poverty of the Linnaean Hierarchy: A Philosophical Study of Biological Taxonomy*. Cambridge: Cambridge University Press.
- (13) Freeman, S. (2008): *Biological Science*. San Francisco: Pearson.
- (14) Gentner, D. (1983): *Structure-Mapping: A Theoretical Framework for Analogy*. *Cognitive Science*, 7, 155-170.
- (15) Goldman, A. (1986): *Epistemology and Cognition*. Cambridge, Massachusetts, London: Harvard University Press.
- (16) Goldstone, R. L.; Medin, D. L.; Gentner, D. (1991): *Relational Similarity and the Nonindependence of Features in Similarity Judgements*. *Cognitive Psychology*, 23, 222-262.
- (17) Goodman, N. (1966 [1953]): *The Structure of Appearance*. Dordrecht: Boston Studies in the Philosophy and History of Science, Synthese Library.
- (18) Goodman, N. (1972): *Problems and projects*. Indianapolis: Bobbs-Merrill.
- (19) Hagen, M. A. (1986): *Varieties of Realism. Geometries of Representational Art*. Cambridge: Cambridge University Press.
- (20) Hahn, U.; Chater, N (1997): *Concepts and Similarity*. En Koen Lamberts & David Shanks (ed.) *Knowledge, Concepts and Categories* (pp. 43-92). UK: Psychology Press.
- (21) Hahn, U.; Chater, N.; Richardson, L. B. (2003): *Similarity as Transformation*. *Cognition*, 87, 1-32.
- (22) Hume, D. (1977 [1739]): *Tratado de la Naturaleza Humana* (Vol. I). (F. Duque, Trad.) Madrid: Nacional.
- (23) Klein, F. (2004 [1924]): *Elementary Mathematics from an Advanced Standpoint. Geometry*. New York: Dover Publications.
- (24) Leitgeb, H. (2007): *A new analysis of quasianalysis*. *Journal of Philosophical Logic*, 36, 181-236.
- (25) Markman, A. B.; Gentner, D. (1993): *Splitting the Differences: A Structural Alignment view of Similarity*. *Journal of Memory and Language*, 32, 517-535.
- (26) Medin, D. L.; Goldstone, R. L.; Gentner, D. (1993): *Respects for Similarity*. *Psychological Review*, 100, 254-278.
- (27) Modenov, P.S.; Parkhomenko, A.S. (1965): *Geometric Transformations. Volume I: Euclidean and Affine Transformations*. New York & London: Academic Press.
- (28) Mormann, T. (1994): *A representational reconstruction of Carnap's Quasi-analysis*. En M. Forbes, (ed.), *PSA 1994*, vol. 1, 96-104.

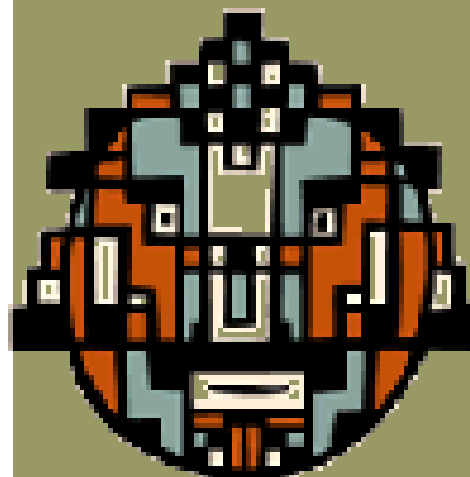


PRIMER PREMIO: Simetría y similitud: la concepción transformacional de la similitud

- (29) Mormann, T. (2009): New Work for Carnap's Quasi-Analysis. *Journal of Philosophical Logic*, 38, 249-282.
- (30) Niiniluoto, I. (1987): *Truthlikeness*. Dordrecht: D. Reidel Publishing Company.
- (31) Nosofsky, R. M. (1991): Stimulus bias, Assymmetric Similarity and Classification. *Cognitive Psychology*, 23, 94-140.
- (32) Nosofsky, R. M. (2011): The generalized context model: an exemplar model of classification. En Emmanuel Pothos (ed.) *Formal Approaches in Categorization* (pp. 18-39). Cambridge: Cambridge University Press.
- (33) Palmer, S. E. (1983): The Psychology of Perceptual Organization: A Transformational Approach. En J. Beck, B. Hope & A. Rosenfeld (eds.), *Human and machine vision* (pp. 269-339). New York: Academic Press.
- (34) Patterson, C. (1982): Morphological characters and homology. En K. A. Joysey & A. E. Friday (ed.), *Problems of Phylogenetic Reconstruction*, 21, (pp. 21-74). London & New York: Academic Press.
- (35) Rosch, E. (1988): Principles of Categorization. En Allan Collins & Edward Smith (ed.) *Readings in Cognitive Science: A Perspective from Psychology and Artificial Intelligence* (pp. 312-322). California: Morgan Kaufmann Publishers.
- (36) Quine, W.V.O. (1970): Natural Kinds. En E. Sosa & J. Kim (eds.) *Metaphysics, An Anthology*. Massachussets: Blackwell Publishers.
- (37) Schreider, J. A. (1975): *Equality, resemblance and order*. Traducido del ruso por M. Greendlinger. Moscow: Mir Publishers.
- (38) Suppes, P.; Krantz, D.; Luce, D.; Tversky, A. (1989): *Foundations of measurement*. New York: Academic.
- (39) Swoyer, C. (1991): Structural representations and surrogative reasoning. *Synthese*, 87, 449-508.
- (40) Tversky, A. (2004 [1977]): Features of similarity. En Eldar Shafir (ed.) *Preference, Belief and Similarity, Selected writings of Amos Tversky* (pp. 7-46). Cambridge & London: MIT press

www.solofici.org

SOCIEDAD DE
LÓGICA,
METODOLOGÍA Y
FILOSOFÍA DE LA
CIENCIA EN
ESPAÑA



SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

Marta Serrapio Lavandeira

Tutor: Carlos Solís Santos

Máster Universitario en Filosofía Teórica y Práctica-UNED

RESUMEN

El objetivo de este trabajo es ofrecer una explicación de la importancia que tuvo el experimento presentado por Newton en 1672 como su *experimentum crucis* para probar su teoría de la luz y los colores y rebatir las teorías que concebían la luz como una modificación del medio. Para entender ese rol “crucial” asignado al experimento en cuestión, empezaremos examinando el contexto teórico a partir del que surge: veremos en qué consistía la nueva metodología experimental de Newton; cuáles eran los rasgos principales de las teorías rivales modificacionistas; en qué se basaba la concepción corpuscular de la luz de Newton; y aclararemos algunas cuestiones relacionadas con las ideas que se tenían acerca de la refracción y los colores antes de que Newton presentase su nueva teoría. A continuación analizaremos la presentación inicial que hizo Newton de su *experimentum crucis* en el artículo de 1672 y, tras esto, exploraremos las principales controversias que este experimento suscitó entre sus principales rivales, que se negaban a aceptar las propiedades de la luz establecidas por Newton: estudiaremos cuáles fueron las críticas de Hooke, Pardies, Huygens, Linus y Lucas. Por último, veremos cómo la publicación posterior de la *Óptica*, junto a algunas demostraciones de los experimentos de Newton realizadas por Desaguliers, jugaron un papel importante para la aceptación de la teoría de Newton. Concluiremos el trabajo considerando las respectivas opiniones de Duhem y Cassini sobre si son posibles los experimentos cruciales en ciencia, para reivindicar, como Cassini, tanto la existencia como la importancia de estos experimentos.



ÍNDICE

1. Introducción: La nueva óptica en el siglo XVII.....	4
2. La metodología experimental de Newton y el rechazo de las hipótesis.....	7
3. Las teorías de la luz como modificación del medio.....	15
4. La teoría corpuscular de la luz de Newton.....	21
5. Concepciones previas sobre la refracción y los colores.....	30
6. La presentación del <i>experimentum crucis</i> en el artículo de 1672.....	37
7. La rivalidad con Hooke y la lucha por la autoridad.....	46
8. La crítica de Pardies.....	51
9. Las reservas de Huygens ante la teoría de Newton.....	59
10. El intento de repetición fallido de Lucas.....	65
11. La aceptación gradual de la teoría de Newton.....	72
12. Conclusión. Consideraciones desde la filosofía de la ciencia posterior sobre los experimentos cruciales.....	81
13. Bibliografía.....	87



SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

I. INTRODUCCIÓN: LA NUEVA ÓPTICA EN EL SIGLO XVII

Aunque hay antecedentes del estudio matemático de ciertos fenómenos ópticos desde la Antigüedad, antes del siglo XVII la cuestión de la naturaleza física de la luz era abordada desde una perspectiva dialéctica o especulativa: se debatía si la luz era una sustancia o un accidente, y si emanaba del ojo o del objeto. A partir del siglo XVII, la luz era un aspecto más de la naturaleza que había de ser estudiado en los mismos términos que el resto de fenómenos naturales. En este siglo la óptica recibe un gran impulso por parte de la astronomía, ya que fue en el marco de la observación de los cielos en el que se investigaron ciertos fenómenos ópticos de importancia, como la refracción y la difracción, que revelaban la existencia de ciertas regularidades en el comportamiento de la luz que debían ser explicadas mediante leyes. Kepler había perfeccionado su teoría de las lentes, estudiando sus combinaciones y aberraciones, y aplicando estos conocimientos al perfeccionamiento de los telescopios. Descartes hallaría la ley de senos para la refracción, según la cual la luz tiene diferentes velocidades para diferentes medios. Grimaldi descubriría el fenómeno de la difracción y Bartholin el de la doble refracción de la luz. Hooke investigaría el fenómeno de los colores de láminas finas.

La óptica moderna es, como buena parte de las nuevas ciencias que empiezan a forjarse en la nueva imagen moderna del mundo, una ciencia abierta, con muchas cuestiones básicas sin resolver. Si bien la óptica geométrica ve grandes avances derivados de la combinación de la experimentación y las matemáticas, esto no es suficiente para dar una explicación completa de todos los fenómenos ópticos, por lo que la cuestión central de la óptica física será el problema de dilucidar en qué consiste la naturaleza física de la luz. La óptica del XVII va a ser, por lo tanto, una ciencia abierta respecto a esta cuestión nuclear y los principales modelos explicativos a finales de este siglo, el corpuscular de Newton y el ondulatorio de Huygens, eran programas abiertos, que no podían explicarlo todo.

Durante el siglo XVII, coexisten dos maneras enfrentadas de entender la naturaleza de la luz: por una parte, encontramos las teorías según las cuales la luz es una modificación mecánica del medio, que eran las dominantes; y, por otra, tenemos la concepción corpuscular de Newton, que daba cuenta de la naturaleza de la luz desde una perspectiva anti-mecánica.

Descartes, en su obra *El mundo. Tratado de la luz*, intenta explicar la luz en términos puramente mecánicos, pero su teoría presentará ciertas inconsistencias que harán que otros investigadores traten de superarla. Hooke, Pardies o Huygens serán los principales continuadores de esa concepción mecanicista de la naturaleza de la luz según la cual ésta es algún tipo de modificación del medio. En general, el principal problema que presentan estas teorías sería que, si la luz es una presión o alteración del medio, no se entiende por qué no se transmite en todas las direcciones, sino en línea recta y formando sombras, como observamos.

Newton, influenciado por el Neoplatonismo de Cambridge, según el cual, además de materia inanimada y pasiva hay un espíritu activo, otorga a la luz un papel central en la naturaleza, al considerarla capaz de agitar y calentar la materia pasiva.

¿Acaso los cuerpos y la luz no actúan mutuamente unos sobre otros? Es decir, ¿no actúan los cuerpos sobre la luz al emitirla, reflejarla, refractarla e inflexionarla y, la luz sobre los cuerpos, al calentarlos y provocar en sus partes un movimiento vibratorio que es en lo que consiste el calor? (Newton, 1977, p. 295).

El estudio que hace Newton de las interacciones de la luz con la materia contribuirán al establecimiento de su teoría corpuscular de la materia y a que crea en la existencia de unas fuerzas inmateriales que actúan a distancia y se asientan en las pequeñas partículas de los cuerpos, siendo las responsables de los fenómenos naturales. De este modo, en su concepción del universo, la luz ocupa un lugar primordial, como podemos ver en este fragmento de las "Cuestiones" de su *Óptica*.

Así pues, viendo que la diversidad de movimiento que encontramos en el mundo está disminuyendo siempre, se presenta la necesidad de conservarlo y reclutarlo mediante principios activos como la causa de la gravedad, por la que planetas y cometas conservan sus movimientos en sus órbitas y los cuerpos adquieren gran movimiento en la caída; o como la causa de la fermentación, por la que el corazón y la sangre de los animales se conserva siempre en movimiento y con calor; las partes internas de la Tierra están constantemente templadas, tornándose muy calientes en algunos lugares; los cuerpos arden y emiten luz, las montañas se incendian, las cavernas de la tierra estallan y el Sol continúa estando violentamente caliente y luminoso, calentando todas las cosas con su luz. De no ser por estos principios, los cuerpos de la Tierra, de los planetas, de los cometas, del Sol y de todas las cosas que en ellos se encuentran se enfriarían y congelarían, tornándose masas inactivas (*ibid.*, p. 344-345).

En definitiva, en la óptica del siglo XVII había acuerdos sobre algunos hechos experimentales, pero desacuerdos sobre las teorías que pudiesen explicarlos. Por tanto, mientras que la explicación de los fenómenos mediante leyes matemáticas era necesaria para que una teoría fuese aceptada, esto no era suficiente. Por ejemplo, en el caso de la ley de senos para la refracción, ésta podía ser derivada correctamente desde diferentes principios físicos que se oponían entre sí, con lo que las matemáticas solas no bastaban. De ahí que surgiesen diferentes teorías explicativas tanto para la óptica geométrica como para la óptica física y de ahí que Newton advirtiese de la importancia de no mezclar las hipótesis no demostradas con las teorías demostradas para el progreso de esta ciencia.

SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

Respecto al fenómeno de los colores, el estudio del mismo cobrará especial relevancia en Newton, puesto que tratará de ofrecer una explicación matemática de los mismos asociándolos con los diferentes índices de refrangibilidad de los rayos de luz y sentando así las bases para el lanzamiento de su nueva teoría de la luz y los colores, según la cual la luz tendría una naturaleza heterogénea. Será un *experimentum crucis* la pieza clave que decidirá a favor de la teoría de Newton y en contra de las teorías de la modificación del medio. Para estas últimas, la luz blanca era un estado simple de la luz, que si se alteraba, por ejemplo, mediante la dispersión a través de un prisma, daba lugar a los colores. Si esto era así, se podía suponer que esas modificaciones mecánicas podrían invertirse. El *experimentum crucis* de Newton demostraba que esto no sucedía, probando así su teoría.

Para entender mejor en qué consistía este experimento, cómo se presentó y cuáles fueron sus implicaciones, debemos sumergirnos antes en el contexto teórico a partir del cual cobraba sentido. Para esto, empezaremos presentando la oposición existente entre Newton y sus rivales: primero veremos cómo la metodología científica que defiende Newton se oponía tanto al método hipotético-deductivo cartesiano como al método baconiano de los naturalistas; a continuación presentaremos los rasgos más salientables de las teorías a las que se oponía Newton, las concepciones de la luz como modificación mecánica del medio; luego introduciremos la concepción corpuscular de la luz de Newton, en concordancia con su teoría atómica de la materia; por último, haremos algunas anotaciones acerca de los estudios sobre la refracción y los colores previos a Newton. Una vez bosquejado este contexto teórico, pasaremos a describir la presentación que hace Newton de su nueva teoría sobre la luz y los colores en 1672, analizando el papel crucial de su *experimentum crucis* para la demostración de la veracidad de su teoría. En lo que sigue, estudiaremos las principales controversias suscitadas por este *experimentum crucis*, examinando las reacciones de los principales rivales de Newton: Hooke, Pardies y Huygens. También veremos las críticas realizadas por los jesuitas Linus y Lucas a su *experimentum*. Por último, esbozaremos cómo la teoría de Newton fue ganando gradualmente aceptación y cómo la publicación de su *Óptica* jugó un papel clave para esa favorable recepción. Concluiremos el trabajo realizando algunas consideraciones desde el punto de vista de la filosofía de la ciencia posterior, respecto a las concepciones de Duhem y Cassini sobre los experimentos cruciales en ciencia.

2. LA METODOLOGÍA EXPERIMENTAL DE NEWTON Y EL RECHAZO DE LAS HIPÓTESIS

Newton creía que la ciencia de su tiempo estaba llena de conjeturas y elementos probabilísticos, pero no ciertamente demostrados, por lo cual veía necesario emprender la tarea de establecer un nuevo método científico que conllevara una ciencia purgada de esos elementos hipotéticos y que se basase en la experimentación para poder establecer modelos matemáticos y físicos que diesen cuenta de los datos observados. Pero la importancia y novedad de esta metodología científica solo se comprenden bien si la comparamos con la metodología cartesiana dominante.

El método científico de Newton, por lo tanto, se opone al método hipotético-deductivo cartesiano, según el cual se argumenta a partir de unas hipótesis o principios *a priori*. Este método fue explicado por Descartes en su *Discurso del Método*, que era una especie de prefacio a los tratados que le seguían: *La Dióptrica*, en el que expone sus concepciones acerca de las propiedades y efectos de la luz¹; *Los Meteoros*, en el que aborda la naturaleza de los cuerpos; y *La Geometría*, en el que trata problemas de geometría y aritmética. Por ejemplo, al inicio de *Los Meteoros*, tras enumerar las cuestiones que pretende explicar en ese tratado, Descartes (1986) advierte de que su argumentación va a ser realizada a partir de la asunción de hipótesis no demostradas:

Verdad es que dependiendo el conocimiento de estas cuestiones de los principios generales de la Naturaleza, que aún no han sido adecuadamente explicados (que yo sepa), será preciso servirme inicialmente de algunas suposiciones, tal y como ya he realizado en la Dióptrica; intentaré que sean tan simples y fáciles que no tengáis apenas dificultades en creerlas, aunque no las haya demostrado (p. 180).

Newton rechazará que las hipótesis sean usadas para determinar las propiedades de las cosas y propugnará que nunca han de ser defendidas en contra de lo que muestran los experimentos. Ese rechazo hacia las hipótesis se debe al uso ilegítimo que hacen de ellas los practicantes del método hipotético-deductivo, para quienes la aceptación inicial de unas hipótesis o principios asumidos dogmáticamente determina la aceptación de lo que muestran los experimentos. Esto no quiere decir que Newton no utilizase hipótesis, sino que el uso que hacía de ellas era completamente diferente, pues éstas no interferían en la aceptación de las propiedades de las cosas que se derivaban de los experimentos.

1. Aunque al inicio de la *Dióptrica* Descartes advierte de que su intención no es abordar el tema de la naturaleza de la luz (cuestión que entre los escolásticos suscitó gran polémica, debatiendo si habría de ser sustancia o accidente), lo cierto es que mediante la descripción de sus propiedades a través de analogías, Descartes nos ofrece pistas de cuál es su concepción de lo que es la luz.

SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

Respecto a sus teorías ópticas, Sabra (1981, p. 298) establece que hay dos niveles explicativos en Newton. Las propiedades de la luz demostradas experimentalmente constituyen teorías que pertenecerían a un primer nivel explicativo: entre éstas, tenemos la teoría de la heterogeneidad original de la luz, la explicación de la ley de refracción por la existencia de una fuerza que actúa perpendicularmente a la superficie refractante y la explicación de los colores de láminas finas en términos de ciertos estados de los rayos que denomina “accesos” de fácil reflexión y fácil transmisión. En un segundo nivel encontraríamos las hipótesis, explicaciones cuya verdad o falsedad es irrelevante para la verdad de las teorías del primer nivel y cuyo objetivo es intentar mostrar a qué podrían deberse esas propiedades establecidas por los experimentos². Por ejemplo, respecto a sus creencias religiosas, Newton las tiene muy en cuenta a la hora de elaborar sus teorías, pero esto no le impide basarlas en el estudio de la naturaleza misma, siguiendo el método inductivo que él propone y rechazando la interferencia de las conjeturas en la aceptación de lo demostrado experimentalmente. De este modo, la adopción de una hipótesis u otra en Newton siempre tenía una base experimental, algo que no ocurría con la aplicación del método hipotético-deductivo. Un ejemplo de esto lo tenemos en la concepción de las interacciones entre la materia, respecto a la cual variaría de postura a lo largo de su vida. Tras haber postulado en su madurez una ontología de fuerzas inmateriales que actúan en el vacío, llegada su vejez admite una nueva hipótesis según la cual dichas interacciones se explicarían debido a la acción de un medio fluido. Así se lo hizo creer un experimento realizado por Desaguliers en 1716, aunque posiblemente sugerido por el propio Newton. El experimento, con dos termómetros que reaccionan de manera similar al calor estando uno en el vacío, parecía mostrar la existencia de un fluido etéreo, sutil y elástico capaz de atravesar los cuerpos. Tras constatar estas observaciones, Newton lanza las posibles hipótesis explicativas de las mismas.

Si se suspenden dos pequeños termómetros en dos grandes recipientes cilíndricos de cristal altos e invertidos, de manera que no toquen los recipientes, y tras ello se saca el aire de uno de esos recipientes y se trasladan ambos de un lugar frío a otro caliente, entonces el termómetro que está *in vacuo* se calentará tanto y casi tan pronto como el que no lo está. Y si se trasladan de nuevo al lugar frío, el termómetro *in vacuo* se enfriará casi tan rápido como el otro. ¿Acaso el calor de la habitación templada no se transmite a través del vacío por las vibraciones de un medio mucho más sutil que el aire y que permanece en el vacío una vez eliminado el aire?. ¿Acaso no es este medio el mismo que aquél en el que la luz se refracta y refleja y por cuyas vibraciones la luz comunica calor a los cuerpos y se pone en accesos de fácil reflexión y fácil transmisión? (Newton, 1977, p. 303).

2. Nótese que esta concepción de las hipótesis es radicalmente diferente a la de los defensores del método hipotético-deductivo.

Es en este sentido en el que debemos entender esa distinción que propone Sabra entre teorías (demostradas) de un primer nivel explicativo e hipótesis de un segundo nivel, que además de tener una base experimental, eran útiles para guiar la investigación científica y para sugerir nuevos experimentos. De este modo, en las “Cuestiones” de la *Óptica*, Newton nos confiesa las hipótesis que tenía en mente y, pese a que sigue siendo cuidadoso con no mezclarlas con las proposiciones que él considera demostradas, nos deja entrever la importancia fundamental que estas hipótesis tenían para él a la hora de disponerse a realizar tales o cuales experimentos. El hecho de que estas “Cuestiones” vayan dedicadas a los futuros investigadores que sucederían al propio Newton con el objeto de guiar sus experimentos, no hace más que reafirmar la relevancia que estas conjeturas tenían para él. Dichas especulaciones arrojaban luz a los oscuros datos y ayudaban a establecer los experimentos adecuados para progresar en el descubrimiento científico. Pero entonces, ¿está operando Newton con una teoría ya en mente, previa a la experimentación? Aunque en sus escritos públicos Newton trata de presentar las propiedades de la luz de acuerdo al método inductivo que defendía, es verdad que sus propias hipótesis sobre la naturaleza de la luz determinan en el modo que acabamos de ver su experimentación, pero esto no implica que siga defendiéndolas a capa y espada en contra de lo que puedan mostrar los experimentos. Así, como considera A. E. Shapiro (1979), podemos afirmar que sí: Newton opera con una teoría ya en mente.

Above all NEWTON was a theoretician. He did not, like so many of his contemporaries, operate in a Baconian mode and set about compiling masses of observations hoping that a theory would emerge. He observed and measured with a theory or idea already in mind (p. 126).

Aunque defiende el uso de un método inductivo, Newton se opone al baconianismo de los naturalistas, muy presente en la primera época de la Royal Society, que se había fundado con una ideología que profesaba la atención a las obras de la naturaleza, por ser lo más seguro que podíamos conocer (además de la *Biblia*). Los defensores acérrimos del método baconiano profesaban un escepticismo hacia los matemáticos, debido a que éstos no dudaban en establecer teorías aunque no hubiesen agotado una recolección de datos que los naturalistas consideraban que cuanto más extensa fuese, mejor. Para Newton, la experimentación y extracción de datos a partir de la naturaleza era muy importante, pero también lo era el establecimiento de modelos matemáticos que acomodasen esos datos. Así, en un borrador al artículo que enviaría a la Royal Society en 1672 conteniendo su nueva teoría sobre la luz y los colores, podemos ver esa oposición de Newton a la ideología de los naturalistas, los cuales, según él, nunca esperarían que la ciencia de los colores se pudiese convertir en un estudio matemático exacto, tal y como él pasará a demostrar.



SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

A naturalist would scarce expect to see the science of those [colours] become mathematicall, & yet I dare affirm that there is as much certainty in it as in any other part of Opticks. ffor what I shall tell concerning them is not an Hypoth{esis} but most rigid consequence, not conjectured by barely infer{ring} 'tis thus because not otherwise or because it satisfies all phænomena (the Philosophers universall Topick,) but evinced by the mediation of experiments concluding directly & without any suspicion of doubt (Newton, 1671/2 ["Draft"], p. 462).



El dominio que Newton tenía de la óptica geométrica le llevó a abordar el fenómeno observado tras la dispersión de los colores a través de un prisma desde un punto de vista consecuente con el método científico que defendía. Así, respecto a las propiedades de la luz y los colores que componen su teoría, dirá que no son hipótesis sino “la más rígida consecuencia extraída a partir de experimentos que concluyen directamente y sin lugar a dudas”. Esta fuerte concepción del método inductivo según la cual se puede demostrar ciertamente una teoría a partir de experimentos, defendida por Newton en una primera etapa, se suavizará en su época de madurez, en la que reconocerá de alguna manera los límites de esta metodología.



En esa primera etapa, en el artículo que envía a la Royal Society en 1672 exponiendo su nueva teoría sobre la luz y los colores, Newton se referirá al que llama *experimentum crucis* como el experimento que prueba positivamente su teoría de la heterogeneidad de la luz y, como consecuencia, excluye a las teorías rivales de la luz como modificación del medio: «And so the *true cause* [...] was detected to be *no other*, then that Light consists of Rays differently refrangible [cursivas añadidas]» (Newton, 1671/2 ["New Theory"], p. 3079). Otro ejemplo de esa concepción fuerte del método inductivo propia de esa primera etapa, lo podemos hallar en su escrito “A Serie's of Quere's”, también de 1672, en el que Newton expone cuál es el método adecuado para establecer teorías verdaderas. Ese método no es el de los naturalistas, que conllevaría una enumeración perfecta de todos los modos posibles de explicar los fenómenos, sino que es el de la deducción de las propiedades de las cosas a partir de los experimentos. Así, en el caso de su nueva teoría de la luz y los colores, la verdad de ésta se derivaría positiva y directamente de la experimentación, y no de la eliminación de las teorías contrarias (esto sería más bien la consecuencia de admitir la verdad de esa teoría y no la causa de esa verdad).

I cannot think it effectual for determining truth, to examin the several waies by which Phænomena may be explained, unless where there can be a perfect enumeration of all those waies. You know, the proper Method for inquiring after the properties of things is, to deduce them from Experiments. And I told you, that the Theory, which I propounded, was evinced to me, not by inferring 'tis thus because not otherwise, that is, not by deducing it only from a confutation of contrary suppositions, but by deriving it from Experiments concluding positively and directly. The way therefore to examin it is, by considering, whether the Experiments which I propound do prove those parts of the Theory, to which they are applyed; or by prosecuting other Experiments which the Theory may suggest for its examination (Newton, 1672 ["A Serie's of Quere's"], p. 5004).

Con el paso del tiempo, Newton suavizará en cierto modo su concepción del método inductivo, admitiendo que si bien no podemos predicar una verdad absoluta de ciertas proposiciones a partir de lo que nos muestran los experimentos, sí debemos admitirlas como “aproximadas a la verdad” hasta que podamos pulir esas proposiciones o establecer las excepciones que puedan implicar. En los *Principia* (1687), Newton refleja ya esta postura y la defiende explícitamente en sus “Reglas para filosofar”, al principio del Libro Tercero. Parte de un principio de economía en la naturaleza: «No deben admitirse más causas de las cosas naturales que aquellas que sean verdaderas y suficientes para explicar sus fenómenos» (Newton, 1987, p. 615). En la Regla IV, defiende el uso del método inductivo porque, según él, pese a las limitaciones de éste, que ahora sí reconoce, es el mejor método que podemos seguir en ciencia para evitar caer en hipótesis.



Las proposiciones obtenidas por inducción a partir de los fenómenos, pese a las hipótesis contrarias, han de ser tenidas, en filosofía experimental, por verdaderas exacta o muy aproximadamente, hasta que aparezcan otros fenómenos que las hagan o más exactas o expuestas a excepciones (*ibid.*, p. 618).

En la “Cuestión 31” de la *Óptica*, Newton se refiere al método inductivo utilizado por él como método de “análisis”: se realizan los experimentos y, a partir de ellos, se extraen por inducción las conclusiones generales, admitiendo solo como objeciones a esas conclusiones aquellas que deriven de otros experimentos y rechazando cualquier hipótesis como posible objeción. En esta “Cuestión”, sin embargo, Newton añadirá algunas aclaraciones en una reedición de la obra en 1717, reforzando el carácter pragmático del método inductivo o de “análisis”. Pese a ser éste el mejor método que cabe usar para investigar la naturaleza, los experimentos por inducción no conllevan una demostración de esas conclusiones generales, siendo necesario admitir las debidas excepciones (que no refutaciones) a lo ya lo ya establecido, que puedan ir surgiendo con el tiempo.

SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

Y, aunque los argumentos a partir de observaciones y experimentos por inducción no constituyan una demostración de las conclusiones generales, con todo, es el mejor modo de argumentar que admite la naturaleza de las cosas y ha de considerarse tanto más fuerte cuanto más general sea la inducción. Si de los fenómenos no surge ninguna excepción, las conclusiones pueden proclamarse en general. Pero, si algún tiempo después surgiese alguna excepción de los experimentos, habrán de comenzar a proclamarse con las excepciones pertinentes (Newton, 1977, p. 349).



En conclusión, el método inductivo defendido por Newton comparte con el de Bacon la parte destructiva de éste según la cual hay que terminar con los “ídolos”, anticipaciones dogmáticas o hipótesis que impregnan la filosofía especulativa. El rechazo de las hipótesis por parte de Newton no implica, como hemos visto, que éste

no utilice hipótesis no demostradas en su investigación científica (explicaciones de un segundo nivel, para Sabra), sino que el uso que hace de ellas es completamente diferente y siempre fundamentado en la experimentación.



Además, Newton va más allá del método baconiano en tanto no se conforma con una recolección indefinida de datos esperando a que de ellos emerjan las leyes de la naturaleza, sino que proclama la importancia de armonizar los datos según teorías matemáticas y físicas que expliquen los hechos “brutos” de un modo siempre fundamentado en la experimentación. Partiendo de esa oposición a los naturalistas, si bien en una primera etapa cree que un experimento puede probar directa y positivamente una proposición teórica, en una segunda etapa transitará hacia una concepción más pragmática de su método inductivo, admitiendo que aunque las pruebas mediante este método no sean absolutas, sí es lo mejor que cabe hacer en ciencia. Seguir el método inductivo de Newton hace que las teorías establecidas de este modo puedan ser sometidas a corrección, mientras que las establecidas de acuerdo al método hipotético-deductivo no pueden ser corregidas ya que los principios metafísicos se aceptan dogmáticamente, incluso en contra de lo que puedan mostrar los experimentos.

3. LAS TEORÍAS DE LA LUZ COMO MODIFICACIÓN DEL MEDIO.

Descartes había elaborado la primera gran teoría mecánica acerca de la naturaleza de la luz, que sería reelaborada por sus sucesores. No obstante, esta teoría conllevaba algunos problemas que Newton no dudó en criticar, como analizaremos a continuación.

Descartes creía que las interacciones entre los cuerpos son materiales y mecánicas, se producen por contacto y no a distancia. Esta forma de explicar el movimiento era consecuente con su concepción de la materia como *res extensa*: el vacío no se admitía y un éter rígido y denso llenaba el espacio. La materia circulaba o se movía en vórtices y Dios, que era el creador del mundo, no intervenía en ese movimiento. Esto para Newton conllevaba un peligro de ateísmo, pero ateniéndose a lo que nos muestra la naturaleza, criticará el éter cartesiano por ser incompatible con los movimientos planetarios. En *El mundo. Tratado de la luz*, Descartes se sirve de ese movimiento vorticial para explicar la experiencia de la luz: la luz sería una presión que procedería del astro solar siguiendo una tendencia centrífuga a alejarse de su centro y se transmitiría a través de la materia hasta alcanzar nuestros ojos. «Pues bien, debe saberse que los hombres de este nuevo mundo serán de una naturaleza tal que, cuando sus ojos sean presionados de ese modo, tendrán un sentimiento totalmente parecido al que nosotros tenemos de la luz» (Descartes, 1989, p. 215).

Fiel a su método hipotético-deductivo, para determinar en qué consiste la luz, Descartes parte de una hipótesis o principio establecido *a priori*: «motion is the only power that can be rationally (that is, clearly and distinctly) asserted to exist in nature» (Sabra, 1981, p. 27). Esto está en consonancia con su concepción de la materia como *res extensa*:

Since extension can be conceived without motion, it is supposed that God, who has created the matter of this new world in the first place, has also actually divided it into parts of diverse shapes and magnitudes (which does not mean creating gaps between them) and set them in motion (*ibid.*, p. 51).

Por lo tanto, el único movimiento que Descartes concibe implica el desplazamiento de los cuerpos de un lugar a otro, ocupando sucesivamente cada posición intermedia. Siendo este movimiento la única fuerza fundamental en la naturaleza (responsable de otras fuerzas como pueden ser el calor, el sonido o la gravedad), Descartes cree que la luz ha de ser también un tipo de movimiento y en la *Dióptrica* explica su funcionamiento mediante analogías, pero evade la cuestión de por qué ha de ser la luz una forma de movimiento. Como tal, la luz extiende sus rayos en un instante del mismo modo que la resistencia de los cuerpos se transmite a través del bastón que usa un ciego para detectarlos.

SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

[...] 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 hacia nuestros ojos a través del aire y de los otros cuerpos transparentes, de igual forma que el movimiento o la resistencia de los cuerpos que encuentra este ciego llega a su mano a través del bastón. Tal consideración os impedirá encontrar extraño, en primer lugar, que la luz pueda extender sus rayos en un instante desde el sol hasta nosotros, pues sabéis que la acción que mueve uno de los extremos del bastón debe alcanzar instantáneamente al otro y que así debería suceder, aunque la distancia entre sus extremos fuese mayor que la existente entre la tierra y los cielos. En segundo lugar, no encontraréis extraño que por medio de la misma podamos ver toda clase de colores, así como que estos colores no sean otra cosa en los cuerpos de color, sino las diversas formas en que los mismos reciben y reflejan la luz contra nuestros ojos (Descartes, 1986, p. 61).



Si la luz no fuese más que una presión propagada sin movimiento efectivo, sería incapaz de agitar y calentar los cuerpos que la refractan y reflejan. Si consistiese en una presión propagada instantáneamente a todas las distancias, exigiría la existencia de una fuerza infinita en todo momento y en toda partícula luminosa a fin de generar tal movimiento. Si consistiese en una presión o movimiento propagado instantáneamente o en el tiempo, entonces se doblaría hacia las sombras, pues la presión o movimiento no se puede propagar por un fluido en línea recta, tras un obstáculo que detenga parte del movimiento; por el contrario, se doblará y extenderá en todas direcciones por el medio en reposo que está del otro lado del obstáculo (Newton, 1977, pp. 313-314).

Newton considera que la propagación de la luz en línea recta es un hecho demostrado. Los fundamentos de esto pueden hallarse en los *Principia*, donde Newton demuestra que «la presión no se propaga a través de un fluido según líneas rectas salvo cuando las partículas del fluido están colocadas directamente en tal sentido» (Newton, 1987, p. 573). Al propagarse en línea recta, la luz no rodea los obstáculos (como hace el sonido). Para Newton, esta propiedad de la luz, su transmisión rectilínea, no podía ser explicada si se concebía la luz como una presión, vibración o modificación del medio, mientras que sí se podía dar cuenta de ella si se consideraba que la luz eran corpúsculos.

En los *Meteoros*, Descartes (*ibid.*) volverá a definir la luz como «la acción o el movimiento de una cierta materia muy sutil, cuyas partes debemos imaginar como pequeñas partículas redondas que ruedan en el interior de los poros de los cuerpos terrestres» (p. 249). Si la luz es un movimiento, podríamos suponer que, como tal, ha de tener una duración finita, por pequeña que sea.

Pero entonces, ¿qué implica realmente la teoría cartesiana de la propagación instantánea de la luz? De acuerdo con esta teoría, la luz no sería un movimiento real, sino más bien una “tendencia al movimiento” que se transmitiría al ojo a través del medio. Pero, ¿por qué una tendencia? Otra vez, Descartes justificará su postura debido a los principios *a priori* adoptados: debido a que la materia es extensa, y como tal, inelástica e incompresible, Descartes deduce que la presión que es la luz ha de transmitirse en un instante.

En definitiva, estas teorías de Descartes recibirán numerosas críticas, pero nos interesa señalar aquí la crítica de Newton a la teoría de la propagación instantánea realizada en la “Cuestión 28” de su *Óptica*. Esta teoría presenta problemas como el no poder dar cuenta de la coherencia del campo visual, o de la propagación rectilínea de la luz (algo que Newton daba por probado). Si la luz es una alteración del medio, ésta debería transmitirse en todas las direcciones, con lo cual no se explica por qué se producen las sombras.

Además de no poder dar cuenta de la transmisión rectilínea de la luz, la teoría de Descartes presentaba otras dificultades. Esa tendencia al movimiento seguiría las mismas leyes que el movimiento (con lo cual no queda del todo clara cuál es la diferencia real entre ambas) y, por ello, Descartes investiga los fenómenos de la reflexión y refracción de la luz por medio de analogías con la reflexión y refracción de proyectiles o cuerpos en movimiento. Esto parece ciertamente incoherente con su pensamiento, ya que si la luz no es un movimiento real sino una presión instantánea, no se entiende cómo para explicar la refracción postula que esa presión tiene diferentes velocidades en diferentes medios.

La concepción de Descartes de la luz como presión continua será reelaborada por sus sucesores de diferentes modos, pero todos seguirán compartiendo que la luz consiste en algún tipo de alteración del medio.

www.solofici.org



Hooke pensará que la luz son vibraciones producidas en un medio homogéneo. En las investigaciones ópticas que Hooke recoge en su *Micrografía* (1665), se centra en el estudio del fenómeno de los colores de las láminas finas. A raíz de eso, proporciona una descripción de ese movimiento en el que consiste la luz: ha de tratarse de un movimiento vibratorio, rápido y corto.



SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

Sería muy largo, digo, insertar aquí el progreso discursivo por el que investigué las propiedades del movimiento de la luz, por lo que sólo aportaré los resultados.

En primer lugar, hallé que debía ser extremadamente *rápido* [...].

A continuación debe ser un movimiento *vibratorio*, a favor de lo cual nos suministra un buen argumento el recién mencionado *diamante*, dado que, si el movimiento de las partes no retornase, entonces, tras muchos frotamientos, el diamante se desintegraría y gastaría [...].

Y, en tercer lugar, creo que los ejemplos extraídos del brillo de los diamantes también harán probable que se dé un *movimiento vibratorio cortísimo* [...].

Creo que éstas son las tres propiedades principales que tiene que tener un movimiento para producir ese efecto que, en el objeto, se denomina luz (Hooke, 1989, pp. 246-247).

Ese movimiento vibratorio será definido también por Hooke como “pulsaciones”, que se expanden en líneas rectas y en todas las direcciones con una igual velocidad a través de un medio homogéneo, formando unas esferas o “frentes de pulsación”, que equipara a las olas provocadas por un sólido en la superficie del agua.



En primer lugar, que debe haber un cuerpo *susceptible* de y *receptivo* a este movimiento que merecerá el nombre de transparente. Luego, que las partes de tal cuerpo han de ser *homogéneas* o del mismo tipo. En tercer lugar, que la constitución y movimiento de las partes ha de ser tal que el pulso del cuerpo luminoso se pueda comunicar o propagar a su través hasta la mayor distancia imaginable en el menor tiempo imaginable, aunque no veo razón para afirmar que haya de ser en un instante, pues no conozco ningún experimento u observación que lo pruebe [...]. En cuarto lugar, el movimiento se propaga en todas las direcciones a través de un *medio homogéneo* mediante líneas *directas* o *rectas*, extendidas en todas direcciones como radios desde el centro de una esfera. En quinto lugar, en un *medio homogéneo* este movimiento se propaga en todas direcciones con una *velocidad igual*, y de ahí que necesariamente todo *pulso* o *vibración* del cuerpo luminoso genere una esfera que aumentará continuamente, creciendo a la manera (aunque indefinidamente más rápido) en que las ondas o anillos de la superficie del agua se amplían en círculos cada vez mayores en torno a un punto en el que se sigue necesariamente que todas las partes de estas esferas que ondulan a través del *medio homogéneo* cortan a los rayos en ángulos rectos (*ibid.*, pp. 247-248).

A raíz de esta explicación de la naturaleza de la luz diferente a la de Descartes pero igualmente mecanicista, Hooke se opondrá a la explicación que hacía Descartes de los colores y considerará que sus observaciones acerca de los colores de láminas finas constituyen una refutación de la teoría cartesiana.

La explicación mecanicista de la luz impulsada por Descartes también se verá continuada en Pardies y en Huygens. Pardies comparte con Hooke la idea de que la luz son vibraciones del medio aunque cree que el frente de la pulsación se transmite de manera diferente, y Huygens creará que la luz consiste en ondas que se producen en la materia etérea y que se expanden esféricamente. Huygens defendía la elasticidad de las partículas de éter, que no son rígidas como las cartesianas, de ahí que la transmisión de esa modificación del medio lleve un tiempo y no sea instantánea, como creía Descartes. Para dar cuenta de la propagación rectilínea de la luz, Huygens se apoyará en la noción de “ondas secundarias”.

Newton no solo arremeterá contra la presunta instantaneidad de la presión que es la luz según Descartes, sino que también criticará cualquier hipótesis que presuponga que la luz es una modificación del medio material, se propague ésta instantáneamente o en un tiempo finito. Para él, los fenómenos lumínicos no dependerán de modificaciones, sino de propiedades originales e inmutables de los rayos que componen la luz:

¿Acaso no son erróneas todas las hipótesis que se han inventado hasta ahora para explicar los fenómenos de la luz mediante nuevas modificaciones de los rayos? Realmente, esos fenómenos no dependen de nuevas modificaciones, como se ha supuesto, sino de propiedades originales e inmutables de los rayos (Newton, 1977, p. 313).



www.solofici.org

4. LA TEORÍA CORPUSCULAR DE LA LUZ DE NEWTON.



Para dar cuenta del movimiento y la interacción entre los cuerpos, Newton tuvo en consideración diferentes principios activos a lo largo de su vida, pasando de creer en su juventud que dichas interacciones se debían a la acción de un éter material a creer en una ontología de fuerzas inmatrimales durante su madurez y volviendo a reconsiderar en su vejez la existencia de un éter sutilísimo, no incompatible con las fuerzas inmatrimales. Todos estos principios activos interactúan con la materia de manera no-mecánica, contrariamente a lo que defendían los partidarios del mecanicismo.

SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

Newton considers various sorts of active agents, some conceived as material but tenuous, others as immaterial, all of them, however, conceived as interacting with matter in non-mechanical ways. An important example of such an active agent is universal gravitation; in the *Principia*, the doctrine of universal gravitation is central to Newton's conception of an omnipresent attractive force. If contact between two bodies is a necessary, though perhaps not a sufficient, condition for mechanical action, gravitational action is nonmechanical. It acts, not with respect to the surface areas of bodies as in the Cartesian conception, but rather with respect to their total quantity of matter. (McGuire, 1977, p. 103).

Si bien esa ontología de fuerzas fue criticada por sus adversarios por considerarla una reinstauración de "cualidades ocultas" en la naturaleza, propia del aristotelismo y contra las cuales se oponía la nueva filosofía mecánica, Newton se defiende de estas acusaciones aclarando que esos principios activos constituyen leyes demostradas por los fenómenos:

No considero que estos principios sean cualidades ocultas, supuestamente derivadas de las formas específicas de las cosas, sino que son leyes generales de la naturaleza por la que se forman las cosas mismas y cuya verdad se nos aparece por los fenómenos, aun cuando sus causas aún no hayan sido descubiertas. Estas cualidades son manifiestas y sólo sus causas son ocultas (Newton, 1977, p. 346).

En el caso de las fuerzas inmateriales, Newton empieza a creer en ellas a raíz de intuir un cierto principio de insociabilidad en la materia que explica la inmiscuidad de ciertos fluidos: «In fact, three years later designated the principle of 'sociableness' and 'unsociableness', force replaces the aether in many contexts as an active agent» (McGuire, 1977, p. 104). En las "Cuestiones" de la *Óptica*, Newton dejará clara la importancia de estas fuerzas.

¿No poseen las pequeñas partículas de los cuerpos ciertos poderes, virtudes o fuerzas con los que actúan a distancia no sólo sobre la luz, reflejándola e inflexionándola, sino también unos sobre otros, para producir una gran parte de los fenómenos de la naturaleza? En efecto, es bien sabido que los cuerpos actúan unos sobre otros por las atracciones de la gravedad, magnetismo y electricidad (Newton, 1977, p. 325).

En el universo newtoniano, las fuerzas, como manifestaciones del espíritu, son lo más importante. La materia, por el contrario, importa en el sentido de que es el asiento de esas fuerzas inmateriales a distancia. A este respecto, Newton creía que la materia está compuesta de partículas duras y permanentes que están dotadas de varias propiedades desde el principio de la creación, por lo que los cambios que se producen en la naturaleza no implicarían la creación o anulación de las propiedades de las partículas, sino que han de ser atribuidos a las asociaciones que se producen entre ellas mismas (causadas por las fuerzas inmateriales).



[...] me parece muy probable que Dios haya creado desde el comienzo la materia en forma de partículas sólidas, masivas, duras, impenetrables y móviles, con tales tamaños y figuras, con tales otras propiedades y en una proporción tal al espacio que resulten lo más apropiadas al fin para el que fueron creadas. Estas partículas primitivas, al ser sólidas, son incomparablemente más duras que cualesquiera cuerpos porosos formados a partir de ellas. Tan duras, incluso, como para no gastarse ni romperse nunca en pedazos, pues ningún poder ordinario es capaz de dividir lo que el mismo Dios ha hecho uno en la primera creación. [...] Por consiguiente, puesto que la naturaleza ha de ser perdurable, los cambios de las cosas corpóreas han de ser atribuidos exclusivamente a las diversas separaciones y nuevas asociaciones de los movimientos de estas partículas permanentes, al ser rompibles los cuerpos sólidos, no en el medio de dichas partículas, sino allí donde se juntan, tocándose en unos pocos puntos solamente (*ibid.*, p. 345).



A partir de esas partículas primitivas se forman los cuerpos porosos, que son tales por componerse de átomos y vacío. En contra de la concepción de la materia como un *plenum*, defendida por Descartes, Newton creía en la escasez de la materia. En la *Óptica* Newton nos ofrece una explicación para imaginar cómo se constituye la materia (mediante una estructura jerárquica de partículas y espacio vacío), aunque aclara que esta explicación no es más que una posibilidad y que podrían darse otras.

De ahí que hayamos de aceptar que los cuerpos son mucho más raros y porosos de lo que normalmente se cree. [...]. Ahora bien, podemos concebir que las partículas de los cuerpos estén de tal modo dispuestas entre sí, que los intervalos o espacios vacíos entre ellas sean iguales en magnitud a todas ellas juntas, y que esas partículas consten de otras partículas mucho más pequeñas que posean tanto espacio vacío entre ellas como partículas menores consten de otras mucho menores, todas las cuales sean igual a todos los poros o espacios vacíos que hay entre ellas, y así continuamente hasta alcanzar a las



SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

partículas sólidas que no tengan poros o espacios vacíos entre sí. [...] Se pueden concebir otras modalidades según las cuales los cuerpos puedan ser extraordinariamente porosos, si bien aún no sabemos cuál es en realidad su trama interna (*ibid.*, pp. 235-237).

En el seno de sus investigaciones ópticas, Newton investiga la composición de los cuerpos, que dejan que la luz penetre en su interior por estar formados de materia y vacío. Así, Newton cree que la luz tiene una naturaleza corpuscular, consistiendo en átomos o partículas muy veloces que se mueven por el espacio vacío. Sus investigaciones empíricas le llevan a creer en esta hipótesis como la mejor manera de entender la naturaleza de la luz para dar cuenta de todos los fenómenos observados. Su creencia en que los rayos de luz tenían ciertos “accesos” y “lados”, como veremos a continuación, y en que los colores descompuestos por un prisma no sufrían ulteriores modificaciones al pasar por un segundo prisma, eran razones a favor de su teoría corpuscular de la luz:

Tras examinar por su cuenta el fenómeno que ya había investigado Hooke acerca de cómo se producen los colores de las láminas finas, Newton también cree que esos colores dependen del grosor de los cristales utilizados: las placas de cristal juegan cierto papel en la producción de esos colores. Newton explica la formación de esos colores mediante su teoría de los accesos: asumiendo que los rayos han sido transmitidos por una superficie refractante, éstos tendrán una disposición a ser fácilmente transmitidos o una disposición a ser fácilmente reflejados.

Todo rayo de luz, al pasar por una superficie refractante cualquiera, adquiere una cierta constitución o estado transitorio que, a medida que el rayo avanza, retorna a intervalos iguales y hace que, en cada vuelta, el rayo adquiera la disposición a transmitirse fácilmente en la siguiente superficie refractante y, entre dos retornos, la disposición a reflejarse fácilmente en ella (*ibid.*, p. 244).

Newton relaciona estos accesos de fácil transmisión y de fácil reflexión con el grosor de las láminas finas, pero lo que nos interesa señalar es que esos accesos serían propiedades originales de los rayos, de acuerdo a lo que afirma en la *Óptica*:

La razón por la que las superficies de todos los cuerpos transparentes gruesos reflejan parte de la luz que incide sobre ellos y refractan el resto, estriba en que algunos rayos se encuentran en accesos de fácil reflexión cuando inciden sobre ellas, mientras que otros se encuentran en accesos de fácil transmisión.

[...] De ahí que la luz se encuentre ya en sus accesos de fácil transmisión y reflexión antes de su incidencia sobre los cuerpos transparentes.

Probablemente se vea puesta en semejantes accesos desde el momento en que es emitida por los cuerpos luminosos, continuando con ellos durante todo su trayecto (*ibid.*, pp. 247-248).

Esa interpretación que hace Newton de los accesos como propiedades originales de los rayos le lleva a ver en este fenómeno un apoyo a su concepción corpuscular de la luz. De hecho, Newton creía que este fenómeno constituía un argumento en contra de la hipótesis modificacionista, como pone de manifiesto en sus “Cuestiones”: «Con esa hipótesis resulta igualmente difícil de explicar cómo es que los rayos pueden estar alternativamente en accesos de fácil reflexión y fácil transmisión» (*ibid.*, p. 315), y continúa argumentando en contra de la posibilidad de explicar estas propiedades de los rayos mecánicamente mediante una teoría del medio. Esto implica que Newton cree que la hipótesis adecuada para dar cuenta de este fenómeno es la corpuscular.

Además de los accesos, Newton creía que los rayos tenían “lados” dotados de propiedades originales. Esto lo deduce de las observaciones realizadas sobre el fenómeno de la doble refracción producida por el cristal de Islandia, ya descrito por Bartholin y por Huygens. Es en las “Cuestiones” donde realiza esa deducción:

Si se coloca sobre un libro un trozo de esta piedra cristalina, cada una de las letras vistas a su través aparecerá doble, debido a una doble refracción. Si un haz de luz incide sea perpendicular o sea oblicuamente sobre cualquiera de las superficies de este cristal, se divide en dos haces por esa misma doble refracción. Ambos haces son del mismo color que el incidente y parecen iguales en cantidad de luz, o casi iguales [...].



Uno de los haces atravesará perpendicularmente al cristal, como sería de esperar de acuerdo con las reglas usuales de la óptica, mientras que el otro [...] se apartará de la perpendicular debido a una refracción inusual (*ibid.*, p. 309).

A continuación Newton describe cómo se producen estas refracciones: una de ellas, usual, se produciría de acuerdo a las reglas de la óptica, y la otra, inusual, según otra regla que detalla seguidamente. Lo más importante es que después de realizar un experimento según el cual Newton hace pasar a los rayos refractados por el primer cristal de Islandia a través de un segundo cristal del mismo tipo, colocado tras el primero de manera que las superficies del segundo queden paralelas a las superficies del primero, se observa que en esta segunda refracción los rayos no sufren alteraciones y se siguen refractando del mismo modo (usual o inusual según qué rayo), de lo cual extrae la

SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

conclusión de que esa diferencia entre los rayos a refractarse de uno u otro modo es una propiedad original de ellos. Esto no parece compatible con las teorías modificacionistas por la siguiente razón:

Si la diferencia, lejos de ser original, se derivase de nuevas modificaciones impresas sobre los rayos en el momento de su primera refracción, se alteraría por las nuevas modificaciones de las tres refracciones siguientes, cuando en realidad no sufre ninguna alteración, sino que es constante y tiene el mismo efecto sobre los rayos en todas las refracciones. La refracción inusual, por tanto, se realiza merced a una propiedad original de los rayos (*ibid.*, pp. 310-311).

A continuación, Newton describe otra observación experimental: si cambiamos la disposición del segundo cristal respecto al primero de modo que los planos de la refracción perpendicular del segundo cristal estén en ángulo recto con los planos de refracción perpendicular del primer cristal, los rayos cambian su forma de refractarse: los que se refractaban de forma usual pasan a hacerlo de forma inusual y viceversa. De esto Newton deduce que los rayos de luz han de tener diferentes lados dotados de diversas propiedades originales: «En efecto, uno y el mismo rayo se refracta a veces del modo usual y a veces del modo inusual, según la posición relativa que mantengan sus lados con el cristal» (*ibid.*, p.311). Rechazadas las hipótesis modificacionistas para dar cuenta de este fenómeno, es evidente que aceptar que los rayos tienen lados implica aceptar el carácter corpuscular de los mismos.

Por tanto, como vemos, son las observaciones experimentales realizadas por Newton las que le llevan a otorgar más probabilidad a su propia hipótesis sobre la naturaleza física de la luz. A lo largo del trabajo desarrollaremos cómo el estudio experimental y matemático de los colores le llevará a ver en la no modificabilidad del color y de los índices de refrangibilidad una prueba clave a favor de su teoría y en contra de las teorías de la modificación, que no podían dar cuenta de estas propiedades de la luz. Pero si bien estos experimentos confirmaban su hipótesis corpuscular, había otros que no casaban con ella, como ocurría con sus investigaciones sobre la difracción.

En el caso de la difracción de la luz, Newton intentó dar cuenta de este fenómeno apoyándose en la teoría corpuscular de la luz y en la actuación de unas fuerzas inmateriales sobre la luz, pero los experimentos realizados no le otorgaron la suficiente evidencia para deducir a partir de ellos ninguna ley o proposición demostrada acerca de este fenómeno. Aunque inicialmente tenía pensado dedicar el Libro III de su *Óptica* a un tratamiento completo de la difracción, al final solamente publicó unas cuantas observaciones experimentales y relegó sus conjeturas acerca de las posibles causas de ese fenómeno a las “Cuestiones”, dejando ese tercer libro de la *Óptica* incompleto.

El título de la única parte que compone el Libro III, reza: “Observaciones relativas a las inflexiones de los rayos de luz y a los colores por ellas producidos”. Vemos cómo Newton evita utilizar el término “difracción”, acuñado por Grimaldi, y en su lugar habla de las “inflexiones” de los rayos con carácter general. Esto es significativo de la postura que defiende Newton, ya que indica que denominar el fenómeno observado por Grimaldi como “difracción” acarrea la connotación de que ese fenómeno implicaría un tipo de transmisión de la luz diferente a la que ocurre en la reflexión y en la refracción. Así, Newton trataba unificar la explicación de todos los fenómenos de acuerdo a su propio marco teórico, según el cual la transmisión de la luz era rectilínea. Así describe en la *Óptica* el problema de la difracción:

Según cuenta Grimaldo, si dejamos que un haz de luz penetre en una habitación oscura a través de un agujero pequeñísimo, las sombras de las cosas situadas en esa luz serán mayores de lo que debieran serlo si los rayos pasasen bordeando los cuerpos en línea recta. Estas sombras poseen además tres bordes, bandas o rangos paralelos y de colores junto a ellas. Si se aumenta la anchura del agujero, las bandas se ensanchan y se mezclan unas con otras, con lo que resultan imposibles de percibir. Algunos han achacado estas sombras anchas y sus bandas a la refracción ordinaria del aire, mostrando un insuficiente examen del asunto (*ibid.*, p. 277).



www.solofici.org

El problema que se encontraba Newton en sus experimentos sobre la difracción era que parecía que la luz se doblaba tras los obstáculos, en contra de lo que ocurriría si los bordease siguiendo una trayectoria rectilínea. Las teorías modificacionistas podrían acomodar mejor este fenómeno, debido a que según ellas la luz es un movimiento de un medio fluido que, para Newton, es diferente del movimiento rectilíneo, como mostraba en los *Principia*: «Todo movimiento propagado a través de un fluido diverge de la trayectoria recta en espacios inmóviles» (Newton, 1987, p. 575).

Newton creía que la imposibilidad de las teorías del medio para dar cuenta de la propagación rectilínea de la luz era la mayor debilidad de éstas, por lo que el hecho de que quizás pudiesen dar cuenta de la difracción de la luz no iba a ser un argumento definitivo a favor de su validez. Aún así, pese a no haber podido deducir ninguna ley acerca de la difracción de la luz, Newton sí se atrevió a publicar en las “Cuestiones” sus conjeturas sobre las causas de ese fenómeno, de acuerdo a su concepción corpuscular. En un principio trató de dar cuenta de las franjas de colores que se producían alrededor de las sombras proyectadas por los obstáculos en términos de lo que ocurre con los rayos de luz en la refracción (en el paso de un medio a otro de diferente densidad), ya que ése era el único modo de

SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

separar los colores, pero sus experimentos le mostraron que esto no era posible: como describe en la "Observación I", la anchura inusual de la sombra de un cabello (más ancha de lo que debiera ocurrir si la luz pasase al lado de ese objeto en línea recta), se produce de igual modo tanto si el cabello está en el aire como en el agua, con lo que se comprueba que el medio no afecta (como debería ocurrir si la desviación de los rayos se debiese a algún tipo de refracción). «Por consiguiente, la gran anchura de estas sombras procede de una causa distinta de la refracción del aire» (Newton, 1977, p. 278). Así, tras describir otra serie de observaciones experimentales, concluye esta única parte que compone el Libro III de la *Óptica*. Abandonado el intento de explicar esta peculiar inflexión de la luz mediante la refracción, Newton, que había suspendido su actividad como experimentador, da paso a sus "Cuestiones", para que otros puedan continuar las investigaciones en su línea teórica. Allí sugiere que la explicación de todas las inflexiones de la luz (no solo de la difracción, sino también de la reflexión y refracción), ha de recaer en la existencia de unas fuerzas inmatrimales que actúan a distancia entre los cuerpos y la luz.

Cuestión 1. ¿Acaso los cuerpos no actúan a distancia sobre la luz y, con su acción, doblan los rayos? ¿No es esta acción (*ceteris paribus*) más fuerte a menor distancia?

Cuestión 2. ¿Acaso los rayos que difieren en refrangibilidad no difieren también en reflexibilidad? ¿No se separan entre sí por sus diferentes inflexiones, para producir, tras tal separación, los colores de las tres franjas anteriormente descritas? ¿De qué modo se inflexionan para producir estas franjas?

Cuestión 3. ¿Acaso los rayos de luz no se doblan varias veces hacia adelante y hacia atrás, con un movimiento como el de una anguila, cuando pasan por los bordes y costados de los cuerpos?

Cuestión 4. ¿Acaso los rayos de luz que caen sobre los cuerpos para reflejarse o refractarse no comienzan a doblarse antes de llegar a los cuerpos? ¿Acaso no se reflejan, refractan e inflexionan por uno y el mismo principio, que actúa diversamente en distintas circunstancias? (*ibíd.*, p. 295).

En las "Cuestiones" que siguen, Newton continúa haciendo referencia a ese "mismo principio" o fuerza inmaterial asentada en la materia pasiva, que haría que los cuerpos actuasen sobre la luz al emitirla, reflejarla, refractarla e inflexionarla y que la luz actúe sobre los cuerpos al calentarlos.

Aunque Newton no pudiese ofrecer una explicación de la difracción de acuerdo a su teoría corpuscular, vemos cómo sus esfuerzos se dirigieron a intentar compatibilizar este fenómeno con su hipótesis corpuscular de la luz, pese a que en este caso no lograrse obtener una prueba a favor de dicha hipótesis. El resultado insatisfactorio de la investigación sobre

este fenómeno es una de las razones por las que dicha hipótesis, a pesar de ser muy útil para Newton desde un punto de vista heurístico, no llegó a formar parte de su cuerpo teórico demostrado experimental y matemáticamente.

Pero a pesar del carácter hipotético de su concepción corpuscular de la naturaleza de la luz, Newton presentó de tal modo sus investigaciones sobre la luz y los colores que, aún restringiéndose a lo que él consideraba proposiciones derivadas de los experimentos, sus adversarios percibían que su argumentación conllevaba implícitamente la asunción de una interpretación corpuscular de la luz, la cual se negaban a aceptar. En lo que sigue, trataremos de desarrollar los pormenores de esa argumentación y las críticas suscitadas, centrándonos en el "crucial" papel que juega el llamado *experimentum crucis* para la demostración de las proposiciones que Newton pretende establecer como probadas y analizando los supuestos implícitos en tal argumentación. Aunque Newton no establezca como probada la corporeidad de la luz, sí va a establecer algunas proposiciones no exentas de polémica debido a que se opondrán a ciertas concepciones aceptadas tradicionalmente sobre la luz y los colores.

5. CONCEPCIONES PREVIAS SOBRE LA REFRACCIÓN Y LOS COLORES.

El fenómeno de la refracción de la luz, esto es, el cambio de dirección que se produce en los rayos de luz cuando pasan de un medio a otro de distinta densidad, fue estudiado seriamente por Ptolomeo en el siglo II, para quien el grado de desviación de los rayos de luz dependía de la diferencia entre la densidad de los medios. Nueve siglos más tarde, Ibn Al-Haytham ofrece una explicación matemática de ese fenómeno, explicando por qué la luz es refractada cuando choca con un medio más denso. Según este autor, la velocidad de la luz disminuye al penetrar en un medio más denso, porque éste ofrecería más resistencia al movimiento de la luz.

When, therefore, light strikes the surface of a denser body, its movement must be altered (*transmutatur*). But since the resistance is not strong enough to repel the movement altogether (as is the case in reflection), the movement is only weakened (Sabra, 1981, p. 95).

Los estudios sobre la refracción de Ibn Al-Haytham tendrían gran repercusión en los investigadores posteriores, como R. Bacon, Witelo y Kepler. Incluso Descartes habría basado gran parte de sus explicaciones de la refracción en Ibn Al-Haytham, pero a diferencia de sus predecesores, Descartes fue el primero en publicar la



SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

ley de la refracción³ y lo hizo en su *Dióptrica*, en 1637. Lo curioso, en este caso, es que Descartes no descubrió la ley “por experimentos”, sino que primero llegó a ella y luego realizó un experimento para verificarla. «He himself wrote in 1632 that the only experiment he had ever made was with the glass designed by Mydorge in 1626 or 1627 for the purpose of verifying the law he had already arrived at» (*ibid.*, p. 103). En la *Dióptrica*, el razonamiento que hace Descartes para defender la ley de refracción se basa en una comparación con el comportamiento de una pelota al chocar contra un medio de diferente densidad: el balón pierde más velocidad al chocar contra un medio blando que contra uno duro. De ahí que, según Descartes, la velocidad de la luz haya de ser mayor en un medio más denso. Aunque no llega a ofrecer una explicación de por qué la velocidad de la luz ha de ser mayor en el agua o en el cristal que en el aire, esto está de acuerdo con los principios *a priori* implicados por su concepción física de la materia y su concepción de la luz como movimiento (instantáneo) de esa materia. Así lo expone en su *Dióptrica*:

Esto dejaréis de juzgarlo extraño, si os acordáis de la naturaleza que he atribuido a la luz, cuando he afirmado que no era sino un cierto movimiento o una acción recibida en una materia muy sutil, que llenaba los poros de los otros cuerpos. Asimismo, si consideraréis que una pelota pierde una mayor cantidad de su agitación chocando contra un cuerpo blando que contra uno que sea duro y que rueda menos fácilmente sobre una alfombra que sobre una tabla totalmente lisa, comprenderéis que de igual forma la acción de esta materia sutil puede verse mucho más impedida por las partes del aire, que, siendo como blandas y no conexas entre sí, no le oponen una gran resistencia, que por las partes del agua, que sí le oponen una mayor resistencia; a su vez, la opuesta por las del agua será mayor que la opuesta por las del vidrio o las del cristal. De suerte que cuanto más duras y firmes son las partículas integrantes de un cuerpo transparente, tanto más fácilmente permiten que la luz los traspase, pues ésta no debe desplazar alguna de sus partes tal y como una pelota debe desplazar las del agua para poder atravesarlas (Descartes, 1986, p. 76).



Hooke difiere de Descartes en el tipo de acción o movimiento que constituye la luz, oponiéndose, como hemos visto, a la teoría de la propagación instantánea e introduciendo la idea de que la luz consiste en ciertos pulsos de un medio homogéneo propagados a gran velocidad. Respecto al fenómeno de la refracción de la luz, Hooke toma como garantizada la fórmula de Descartes e intenta explicar este fenómeno dando cuenta de lo que le sucede a un frente de pulsación cuando pasa de un medio a otro. Aquí Hooke reemplaza la explicación mediante comparaciones que ofrecía Descartes por una explicación mecánica pero, al igual que él, acaba deduciendo que la velocidad de la luz debe de ser mayor en los medios más densos.

www.solofici.org

Por otra parte, Fermat también criticará la explicación que Descartes ofrece para la ley de refracción y, aunque llega a la misma ecuación que él⁴, afirmará que la velocidad de la luz es mayor en los medios menos densos, siguiendo el principio del menor tiempo⁵. Pardies intenta hacer una demostración de ese principio aplicado a la refracción, lo que le lleva a suscribir la ley de refracción de Fermat y la consideración de que la velocidad de la luz es menor en los medios más densos. Huygens sigue la línea de Fermat y Pardies en relación a la refracción, y se opone a la teoría de la propagación instantánea de Descartes considerando que la luz no es una tendencia al movimiento sino un movimiento real de un medio que es elástico: la luz consiste en ondas de la materia etérea elástica. Huygens admite la existencia del vacío y cree que las partículas de éter deben ser capaces de contraerse y expandirse, lo cual requiere un tiempo. Para él, la propagación de la luz es rectilínea y la explica mediante el concepto de ondas secundarias (en la superficie de cada onda, cada punto ha de ser visto como el centro de una onda secundaria que se mueve a la misma velocidad que la onda principal) que permanecerían perpendiculares a la dirección de propagación tras la refracción. En su obra *Traité de la Lumière* (1690), trata de mostrar cómo el principio del menor tiempo puede ser deducido a partir de su ley de ondas secundarias.

3. La ley es expresada por las ecuaciones: $\sin i = n \sin r$, donde n es una constante, i se refiere al ángulo de incidencia y r al ángulo de refracción; y $v_r = nv_i$, donde v_r es la velocidad de refracción y v_i la velocidad de incidencia.

En la misma época, Snell también llega a la ley de refracción. Ha habido cierta polémica respecto a si Descartes plagió la ley de Snell, ya que parece que pudo estar al tanto de los escritos de Snell, pero debido a que no hay evidencia suficiente para defender esta postura, asumimos, como Sabra (1981, pp. 99-105), que el descubrimiento de Descartes fue realizado de manera independiente.

4. Decimos que Fermat llega a la misma ecuación que Descartes, pero ofrece las velocidades en una *ratio* inversa a los senos. Pardies y Huygens defenderán esta presentación de la ley, frente a la propuesta por Descartes y apoyada por Hooke. Cf.: Sabra, 1981, p. 197.

5. Según este principio, la naturaleza siempre actúa por las trayectorias que requieren el menor tiempo. Clerselier critica este principio por considerarlo metafísico: defenderlo es atribuir elección y conocimiento a la naturaleza, y él cree que la naturaleza actúa por determinismo.

SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

This law implies that when the angle of refraction is smaller than the corresponding angle of incidence, the velocity must have been diminished by refraction. And since light in passing from a rare into a dense medium is deflected towards the normal, it must be concluded that the velocity of light is greater in rarer media (Sabra, 1981, p. 218).

Respecto a la ley de refracción, Newton afirmará las ecuaciones propuestas por Descartes pero les dará una interpretación dinámica completamente distinta a la de Descartes, ya que éste pretendía construir toda su física sobre la idea de movimiento, excluyendo cualquier otra fuerza. De acuerdo a su segunda ley del movimiento, Newton explica la refracción diciendo que hay una fuerza actuando perpendicularmente en la superficie del medio refractante. De acuerdo con las bases de su física, Newton cree que los rayos siguen las leyes de la dinámica, de modo que esa fuerza perpendicular sería la causante de la refracción y de su actuación se deriva la proposición de que la velocidad de la luz es mayor en los medios más densos. Esto coincide con lo que decían Descartes y Hooke, pero a diferencia de ellos, Newton trata de demostrar que dicho fenómeno se produce debido a la existencia de esa fuerza actuante. Esa suposición distaría, según él, de ser considerada una hipótesis, ya que no determina en qué consiste la naturaleza de la luz o por qué tipo de fuerza es refractada. En la "Cuestión 29" de su *Óptica*, Newton explicita que esa fuerza que actuaría a distancia se parecería a un tipo de fuerza atractiva.

Las sustancias transparentes actúan a distancia sobre los rayos de luz al refractarlos, reflejarlos e inflexionarlos, a la vez que, por su parte, los rayos de luz agitan a distancia las partes de esas sustancias para calentarlas. Además, esta acción y reacción a distancia se asemeja muchísimo a una fuerza atractiva entre los cuerpos. Si la refracción se debe a la atracción de los rayos, los senos de incidencia deben de estar con los senos de refracción en una proporción determinada, como hemos demostrado en nuestros Principios de Filosofía, siendo esta regla verdadera por experiencia (Newton, 1977, p. 321).



Pasando a la cuestión de los colores, Descartes los descomponía mediante el uso de un prisma y explicaba su formación atribuyendo una velocidad rotatoria diferente a los glóbulos de la materia correspondientes, aunque no ofrecía ningún medio para medir dicha velocidad. Cuando esos glóbulos limitan por un lado con una parte del medio en sombra (en reposo) y por otro lado con una parte del medio en luz (en movimiento), es cuando su velocidad rotatoria varía y se producen los colores. Esto lo explica Descartes (1986) en los *Meteoros*, donde, no obstante, también admite que no siempre es necesario la limitación entre la luz y la sombra para esa producción de los colores:

Admito, no obstante, que la sombra y la refracción no son siempre necesarios para producirlos, y que, en su lugar, el grosor, la figura, la situación y el movimiento de las partículas de los cuerpos que se llaman coloreados, pueden concurrir de modo diverso junto con la luz para aumentar o disminuir el giro de las partículas de la materia sutil (Descartes, 1986, p. 252).



Hooke utilizará por primera vez el término *experimentum crucis* para referirse a las observaciones que él lleva a cabo en su *Micrografía* respecto a los colores de las láminas finas, y que refutarían la doctrina cartesiana de la formación de los colores, según la cual los colores se deben a esa rotación de los glóbulos que constituyen el medio etéreo y a su limitación entre la luz y la sombra. Dice Hooke:

[...] no puedo imaginar cómo se pueda dar una razón plausible, mediante la *hipótesis cartesiana*, de la naturaleza de los colores engendrados en las láminas delgadas de estas nuestras *observaciones microscópicas*, pues en muchas de éstas, las superficies refractantes y reflectantes son paralelas entre sí, con lo que no se puede engendrar ninguna *rotación*, sin que haya tampoco necesidad alguna de una sombra o límite de los rayos brillantes, tal y como se supone necesario [...] para la generación de cualesquiera colores distintos (Hooke, 1989, p. 257).

El *experimentum crucis* de Hooke mostraría que los colores en las láminas finas transparentes son generados por la reflexión, refutando así la teoría cartesiana según la cual para la producción de los colores es necesaria una rotación de los glóbulos y un límite entre luz y sombra. Para poder dar cuenta de este fenómeno, Hooke trata de demostrar primero que tanto la reflexión como la refracción provocan una oblicuidad en el pulso de los rayos respecto a sus líneas de transmisión, y esa propiedad es la que generaría el color.

Así pues, habiendo mostrado que existe tal propiedad en el *prisma* y en el *glóbulo* de agua, mediante la cual el pulso se torna *oblicuo* a la progresión, y tanto más cuanto mayor sea la refracción, consideraré a continuación cómo ello conduce a la producción de colores y qué tipo de impresión causa en el fondo del ojo (*ibid.*, p. 259)



Según la teoría fisiológica de los colores de Hooke, todos los rayos tienen sus pulsos oblicuos a su progresión y cada rayo tiene «dos propiedades o colores; a saber, un *rojo* en un lado y un *azul* en el otro» (*ibid.*, p. 261). Así, dependiendo de la manera en que los rayos entren en el ojo, ya que sufren una refracción al atravesar la córnea, «según que se vea estorbado este o aquel lado o extremo del pulso, variarán las *impresiones* de la *retina*» (*ibid.*).

SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

De la consideración de las propiedades de dichas impresiones podemos derivar estas breves definiciones de los colores: que *el azul es una impresión sobre la retina de un pulso de luz oblicuo y confuso, cuya parte más débil precede y la más fuerte sigue*. Y que *el rojo es una impresión sobre la retina de un pulso de luz oblicuo y confuso cuya parte más fuerte precede y la más débil sigue* (*ibid.*, p. 262).

Para Hooke, los colores primarios son el rojo y el azul. Esta creencia se debe a que el haz de luz que emplea es demasiado grueso y no utiliza una distancia considerable para proyectar los colores separados por refracción, con lo cual solo observa esos dos colores. Newton variará estas circunstancias: aumentará la distancia entre el prisma y la pared en la que observa los colores proyectados, y disminuirá el tamaño del agujero a través del que pasa la luz del sol, de modo que puede observar todo el espectro de colores. No obstante, establecidas por Hooke las propiedades de esos pulsos de luz que constituyen el rojo y el azul, éste propone que esas características del movimiento en el que consiste la luz son también las causantes de la producción de los colores en las láminas finas transparentes:

Se puede ver fácilmente que estas propiedades, que en el caso del prisma y las gotas de lluvia que caen han demostrado ya ser la causa de los colores generados, son también las productoras de los colores que aparecen en los cuerpos *laminados* transparentes, para cuya explicación se ha antepuesto todo esto (*ibid.*).

Para esa aparición de los colores en láminas delgadas, es necesario que haya un cuerpo reflectante en el lado inferior de las mismas. La teoría de Descartes no daba cuenta de la producción de los colores en láminas finas por reflexión.

Newton intentará superar la teoría de la producción de los colores de Hooke, partiendo de una concepción de la luz y los colores radicalmente diferente y opuesta. Para Newton, los colores no son modificaciones que la luz sufre tras la reflexión o refracción, sino propiedades originales e innatas a los rayos, así como lo son sus distintos grados de refrangibilidad. Newton deduce esto de la realización de una serie de experimentos con prismas (entre ellos, su *experimentum crucis*), haciendo aparecer los colores a partir de la refracción de la luz blanca. Newton relacionará directamente a los colores con el índice de refrangibilidad de los rayos, centrandó la explicación de los colores en el problema geométrico de la refracción, lo cual analizaremos más adelante. Lo que nos interesa resaltar aquí es que la nueva concepción que Newton propone de los colores se opone a la concepción de éstos como meras modificaciones de la luz, mantenida por sus rivales, para quienes la luz es algo homogéneo y no un agregado heterogéneo de rayos, como pretende establecer Newton.

6. LA PRESENTACIÓN DEL *EXPERIMENTUM CRUCIS* EN EL ARTÍCULO DE 1672

En relación a sus teorías ópticas, el único artículo publicado por Newton durante el siglo XVII fue su famoso "New theory about light and colors", que envió a la Royal Society en 1672 y salió a la luz en las *Philosophical Transactions* ese mismo año. No obstante, antes de eso, Newton había estado investigando ampliamente ciertos fenómenos ópticos y preparando un tratado llamado *Optical Lectures*, que sin embargo no llegó a publicar en vida.

In it he gave a thorough account of his discovery of the unequal refrangibility of lights of different color, developed his theory of color and the nature of sunlight -both supported by copious experiment- and attempted to develop a mathematical theory of color (Shapiro, 2008, p. 419).

Sin embargo, aunque la publicación de ese tratado sería descartada por Newton, cuando éste decide presentar su teoría a través del breve artículo de 1672, reconoce que llevaba tiempo investigando ciertos fenómenos ópticos y proclama que las investigaciones que le han llevado a esa nueva teoría sobre la luz y los colores se remontan a 1666, su *annus mirabilis*. Westfall (1980) profundiza en los matices de la confusa cronología de las tempranas investigaciones ópticas de Newton y, si bien parece no estar muy claro el orden de los experimentos que realizó, lo que sí parece evidente es que Newton partía de un rechazo a las teorías vigentes sobre el color y la luz blanca⁶. La publicación de la *Micrografía* de Hooke en 1665 le estimula para oponerse a la explicación de la luz como pulsos del medio y de los colores como impresiones causadas por esos pulsos. Pero pese a que Hooke pasaría a ser su enemigo declarado, Newton no solo se oponía a él con su nueva teoría sino a todas las explicaciones alternativas de los colores aparentes, presentes en ese momento.

www.solofici.org

Like other mechanical philosophers, Hooke had merely provided a mechanism for the existing theory of apparent colors, phenomena such as the rainbow and the colored fringes seen through telescopes and prisms. The theory had been fatally easy to mechanize. It had employed a scale of colors, which was also a scale of strength, running from brilliant red, considered to be pure white light with the least admixture of darkness, to dull blue, the last step before black, which was the complete extinction of light by darkness (Westfall, 1980, p. 159).

6. Entre estas teorías se encontraban la de Descartes, la de Boyle (*Experiments and Considerations Touching Colors*, 1664) y la de Hooke (*Micrografía*, 1665). Para obtener más detalles sobre la cronología de los primeros experimentos ópticos realizados por Newton véase: Westfall, 1980, pp. 156-160.

SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

En su "New Theory of Light and Colors" Newton presenta una serie de experimentos que, según él, le llevaron a descubrir y probar que la luz no es homogénea, como se creía, sino que consiste en rayos de diferente refrangibilidad, y que los colores no son cualificaciones de la luz, sino propiedades originales de la misma. Este artículo envuelve una pretensión de refutación de las teorías alternativas para proclamar la suya como la única que puede dar cuenta de los fenómenos observados. Como defiende A. Raftopoulos (1999):

[...] the order of the arguments and experiments presented in the paper is not the historical but the logical order which renders the arguments more clear, stronger and less ambiguous. [...] It follows that Newton's main concern is not to present an accurate picture of the formation of his theory, but to convince his readers of its truth (p. 98).



Teniendo esto en cuenta, el primero de los experimentos que Newton describe consiste en situar un prisma triangular junto a un agujero en la pared de una habitación a oscuras, de modo que la luz del sol que entra por el agujero de la pared es refractada por el prisma hacia la pared opuesta, formándose en ella una imagen que es oblonga en vez de circular, como se supone que debería ser de acuerdo a las leyes de la refracción. El problema consiste en explicar la formación de esa figura oblonga. Para ver a qué podía deberse, Newton examina posibles causas, como el grosor del cristal del prisma, que descarta tras variarlo y ver que no altera la formación del espectro en la pared. También rechaza que los colores refractados se produzcan debido a una mezcla de luz y oscuridad, la cual provendría de las paredes del agujero: para ello prueba a variar el tamaño del agujero y observa que el modo en que aparecen los colores en la pared no cambia.

Otra posible causa examinada por Newton es la posible influencia de irregularidades del cristal del prisma. Para comprobar si efectivamente la figura oblonga del espectro se debe a algún tipo de irregularidad en el prisma, Newton realiza un experimento combinando dos prismas de manera que sus respectivos ejes queden paralelos, uno con el vértice arriba y otro con el vértice abajo: pensaba que si las irregularidades del prisma eran las causantes de esa dispersión de los rayos, un segundo prisma que refractase los rayos que salen del primer prisma debería aumentar los efectos irregulares de la dispersión y eliminar los regulares. Sin embargo, Newton halla que la figura formada tras pasar la luz por el segundo prisma es tan regular como la observada anteriormente.

Otra posibilidad examinada por Newton es que la figura oblonga que forman los rayos en la pared se deba a la diferente incidencia de los rayos provenientes de diferentes partes del sol. Tras una serie de mediciones llega a la conclusión de que es imposible que esa diferencia en la incidencia de los rayos de sol pueda hacer divergir a los rayos en un ángulo tan grande como el observado en la figura oblonga.

A continuación, Newton también estudia la posibilidad de que los rayos viajen en línea curva tras pasar por el prisma, de igual manera a como se movería un balón de tenis tras chocar con una raqueta, dando lugar a la figura oblonga debido a esa trayectoria. Newton parece rebatir aquí la idea de Descartes según la cual la oblicuidad de los rayos se debería a los cambios de velocidad de los glóbulos rotativos cuando pasan de un medio a otro.

For, a circular as well as a progressive motion being communicated to it by that stroak, its parts on that side, where the motions conspire, must press and beat the contiguous Air more violently than on the other, and there excite a reluctancy and reaction of the Air proportionably greater. And for the same reason, if the Rays of light should possibly be globular bodies, and by their oblique passage out of one medium into another acquire a circulating motion, they ought to feel the greater resistance from the ambient Æther, on that side, where the motions conspire, and thence be continually bowed to the other. But notwithstanding this plausible ground of suspition, when I came to examine it, I could observe no such curvity in them. And besides (which was enough for my purpose) I observed, that the difference 'twixt the length of the Image, and diameter of the hole, through which the light was transmitted, was proportionable to their distance (Newton, 1671/2 ["New Theory"], p. 3078).

Tras descartar estas posibles causas de la longitud de la imagen formada en la pared, que no se corresponde con lo que debería ocurrir de acuerdo a las leyes de la refracción admitidas generalmente, Newton se dispone a desafiar la asunción implícita en esas leyes de la refracción, a saber, que los rayos de sol son igualmente refrangibles. Así, Newton centra la atención en el problema mismo de la refracción:

The real problem was one of refraction; it consisted in explaining the *shape* of the spectrum for a given position of the prism, and this on the supposition that the rays from the sun were all equally refrangible. That supposition was *implicitly* accepted by all writers on optics, or, at least, it had not been challenged by anyone. It was on the basis of that supposition that Newton had expected to see the solar image in a round, and not as he found, in an oblong shape. After having dismissed some of the most obvious hypotheses that might be thought to account for the elongation of the image, he performed the *experimentum crucis* (Sabra, 1981, p. 241).



SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

Para desafiar esa supuesta igual refrangibilidad de los rayos, Newton lleva a cabo el que considera su *experimentum crucis*, denominado así por la crucialidad que juega en la argumentación a favor de su teoría: este experimento permitiría decidir entre dos tipos de explicación incompatibles entre sí: su objetivo es refutar las teorías modificacionistas de la luz y probar la nueva teoría de Newton. Tras describir en qué consiste este experimento, Newton se dispone a describir positivamente esa teoría, estableciendo ciertas propiedades sobre la luz y los colores, que considera probadas por los fenómenos. Veamos cómo describe Newton este experimento crucial:

I took two boards, and placed one of them close behind the Prisme at the window, so that the light might pass through a small hole, made in it for the purpose, and fall on the other board, which I placed at about 12 feet distance, having first made a small hole in it also, for some of that Incident light to pass through. Then I placed another Prisme behind this second board, so that the light, trajected through both the boards, might pass through that also, and be again refracted before it arrived at the wall. This done, I took the first Prisme in my hand, and turned it to and fro slowly about its Axis, so much as to make the several parts of the Image, cast on the second board, successively pass through the hole in it, that I might observe to what places on the wall the second Prisme would refract them. And I saw by the variation of those places, that the light, tending to that end of the Image, towards which the refraction of the first Prisme was made, did in the second Prisme suffer a Refraction considerably greater then the light tending to the other end (Newton, 1671/2 ["New Theory"], pp. 3078-3079).

Tras haber hecho pasar sucesivamente los diferentes rayos de colores descompuestos por el primer prisma a través del agujero que lleva al segundo prisma, mediante la rotación del primer prisma, Newton percibe que unos colores son más refractados que otros: los colores que estaban más bajos en la imagen oblonga formada tras pasar la luz por el primer prisma son más refractados que los que estaban más altos en la imagen. A partir de esta observación, Newton cree que se puede afirmar la naturaleza heterogénea de la luz, es decir, que ésta consiste en rayos que tienen diferentes índices de refrangibilidad. Así, establece:

And so the *true cause* of the length of that Image was detected to be *no other*, then that Light consists of Rays differently refrangible, which, without any respect to a difference in their incidence, were, according to their degrees of refrangibility, transmitted towards divers parts of the wall [cursivas añadidas] (*ibid.*, p. 3079).

Como habíamos dicho, Newton concentra esta importante conclusión en el problema de la refracción misma, y no en el de la apariencia de los colores. Con ello ataca de raíz a las concepciones rivales según las cuales la luz es homogénea, estableciendo que la diferente refrangibilidad de los rayos es una característica inherente a la luz. De hecho, si nos fijamos en las palabras marcadas en cursivas en el texto anterior, vemos cómo el modo en que Newton redacta su conclusión implica una aniquilación de las posibles explicaciones alternativas: él nos ofrece la "verdadera causa" del fenómeno observado, por lo que no cabe la posibilidad de otras explicaciones alternativas. Así, más adelante en ese mismo artículo de 1672, Newton niega explícitamente que la luz pueda ser homogénea:

www.solofici.org

[...] Light is not similar, or homogeneal, but consists of difform Rays, some of which are more refrangible than others: So that of those, which are alike incident on the same medium, some shall be more refracted than others, and that not by any virtue of the glass, or other external cause, but from a predisposition, which every particular Ray hath to suffer a particular degree of Refraction (*ibid.*, p. 3081).

El *experimentum crucis* le permite a Newton afirmar su propia teoría tras haber descartado previamente otras posibles explicaciones del fenómeno observado en la pared. Así, tras establecer la novedosa conclusión de que la luz consiste en rayos que tienen una diferente refrangibilidad, Newton despliega su doctrina en una serie de proposiciones que se derivan de esa conclusión y que conciernen a la naturaleza de la luz y los colores. Hablar de la diferente refrangibilidad de los rayos es para Newton hablar de la diferente disposición de los mismos a exhibir un color u otro. De este modo, establece la siguiente proposición en relación a la naturaleza de los colores:



Colours are not *Qualifications of Light*, derived from Refractions, or Reflections of natural Bodies (as 'tis generally believed,) but *Original and connate properties*, which in divers Rays are divers. Some Rays are disposed to exhibit a red colour and no other; some a yellow and no other, some a green and no other, and so of the rest. Nor are there only Rays proper and particular to the more eminent colours, but even to all their intermediate gradations (*ibid.*, p. 3081).

A continuación, Newton afirma que al mismo grado de refrangibilidad siempre corresponde el mismo color, y viceversa: al mismo color siempre corresponde un mismo grado de refrangibilidad. Ese color y grado de refrangibilidad propios de cada rayo no se ven modificados por la refracción, reflexión o cualquier otra causa. Cuando los rayos que componen la luz blanca son separados por refracción (u otra causa), aparecen los colores: Newton aclara que éstos no son generados o creados en

SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

ese momento, sino que simplemente emergen o se hacen visibles gracias a esa separación de los rayos. De este modo, existirían dos tipos de colores: los primarios o simples y los compuestos. Los colores primarios son los del arcoíris (entre los que están el rojo, amarillo, verde, azul, violeta-lila, naranja, índigo y otros colores intermedios), y los compuestos se derivarían de la conjunción de aquéllos. Pero lo más sorprendente, dice Newton, es el hecho de que la luz blanca esté compuesta por todos los colores primarios, de modo que siendo la luz un agregado heterogéneo de rayos con sus correspondientes colores, resulta que el blanco es la suma de todos ellos y el color usual de la luz.



Tras el establecimiento de esta serie de proposiciones tocantes a la luz y los colores, Newton ya puede explicar a qué se debe la forma oblonga de la imagen refractada en la pared: como los rayos son diferentes entre sí y cada uno posee una mayor o menor disposición a ser refractado por el prisma, de acuerdo a un color concreto, cada rayo es refractado en la figura de la pared siguiendo un orden desde el color menos refrangible, el rojo escarlata, al más refrangible, el violeta. Así da Newton por resuelto el problema de que la figura oblonga no se corresponda con lo que debiera ocurrir de acuerdo a la asunción de que los rayos de luz fuesen igualmente refrangibles, una asunción que para Newton ha resultado carecer de fundamento. Los rayos no se refractan todos ellos de la misma manera, por lo que la luz no es homogénea, como se pensaba, sino un agregado heterogéneo, pero ¿conlleva esa afirmación la proclamación de una naturaleza corpórea de la luz? En el artículo de 1672, Newton sí sugiere que la luz tiene una naturaleza corpórea, pero no llega a establecerlo como si se tratase de una proposición probada:

These things being so, it can no longer be disputed, whether there be colours in the dark, nor whether they be the qualities of the objects we see, no nor perhaps, whether Light be a Body. For, since Colours are the *qualities* of Light, having its Rays for their intire and immediate subject, how can we think those Rays *qualities* also, unless one quality may be the subject of and sustain another; which in effect is to call it *Substance*. We should not know Bodies for substances, were it not for their sensible qualities, and the Principal of those being now found due to something else, we have as good reason to believe that to be a Substance also.

Besides, whoever thought any quality to be a *heterogeneous* aggregate, such as Light is discovered to be. But, to determine more absolutely, what Light is, after what manner refracted, and by what modes or actions it produceth in our minds the Phantasms of Colours, is not so easie. And I shall not mingle conjectures with certainties (*Ibid.*, p. 3085).

El problema que expresa Newton aquí consiste en que si consideramos, como él hace, que los colores son cualidades o propiedades de los rayos de luz, ¿cómo vamos a defender que los rayos son también cualidades?, lo cual nos lleva a preguntar ¿cómo va a ser la luz una cualidad? Si la luz o los rayos fuesen cualidades, tendríamos que asumir que los colores son una cualidad de otra cualidad, a saber, del agregado heterogéneo en que consiste la luz. Parece más lógico pensar que una cualidad ha de serlo de una sustancia pero, no obstante, Newton no se adentra en este pantano y se limita a concluir que no es tarea fácil esclarecer en qué consiste la naturaleza de la luz y a retomar uno de sus preceptos metodológicos más importantes, que es el de no mezclar las conjeturas o hipótesis con las certezas o proposiciones demostradas.

En apartados anteriores ya hemos visto que Newton realmente creía en una naturaleza corpuscular de la luz, en concordancia con su concepción corpuscular de la materia, pero era consciente de que esa hipótesis no tenía el mismo carácter que sus teorías o leyes matemáticas y experimentales. No obstante, para Newton las hipótesis de la luz como modificación del medio no estaban en ese mismo nivel que su teoría corpuscular: las hipótesis modificacionistas no podían ser defendidas debido a que eran inconsistentes con la propagación rectilínea de la luz. La defensa de esta hipótesis por parte de sus rivales era apriorística y sostenida en contra de lo que mostraban los fenómenos.



Los rivales de Newton, no obstante, creerán que la argumentación de este conlleva una defensa de la teoría corpuscular, pese a haber advertido que no quería mezclar las conjeturas sobre la naturaleza de la luz con las proposiciones establecidas en su teoría. ¿Tenían argumentos sus rivales para pensar que esa teoría implicaba una concepción corpuscular? Según Newton, no se puede afirmar con propiedad lo que es la luz, si una sustancia o una cualidad, pero lo que sí considera que está demostrado es que la luz es un agregado heterogéneo de rayos que difieren en refrangibilidad. Ahora bien, al indicar que no es posible concebir un agregado heterogéneo como una cualidad, parece obvio que está conduciendo a sus lectores a concluir que efectivamente ese agregado que es la luz no es sino una sustancia. Así lo expresa Sabra (1981):

That white light is a heterogeneous mixture is a thesis which Newton never abandoned; it was for him an experimental fact proved beyond any possible doubt. But since he expressed himself in terms of this dichotomy, substance or quality, was it not only natural for his readers to assume that he favoured a corpuscular hypothesis about light? [...] What was certain for Newton? That light is a heterogeneous aggregate. But if, as he himself argued, it was not possible to conceive of quality as such an aggregate, did he leave much doubt as to the conclusion he wanted to draw? (p. 243).

SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

Tras la publicación de su “New Theory of Light and Colors”, se disparará la alarma entre los rivales de Newton, entre los cuales encontraremos a Hooke, Pardies y Huygens. Todos ellos creen que Newton está defendiendo una concepción corpuscular de la luz, por lo que todos se referirán a esa “nueva teoría” newtoniana como hipótesis, dando a entender que los hechos pueden ser interpretados de otro modo. Todos ellos seguirán la tradición según la cual la luz es concebida como una alteración del medio. A su vez, Newton dirá que esa concepción es tan hipotética como su concepción corpuscular de la luz y negará que esta última forme parte de su teoría demostrada, al ver que se estaban mezclando sus proposiciones y observaciones experimentales con esa hipótesis.

7. LA RIVALIDAD CON HOOKE Y LA LUCHA POR LA AUTORIDAD

Después de que Newton enviase su “New Theory about Light and Colors” a la Royal Society en 1672, Hooke fue elegido por Oldenburg, responsable de la correspondencia entre los miembros de la Royal Society, para redactar sus “consideraciones sobre la teoría de Newton”, convirtiéndose en el primer crítico de esta nueva teoría. Su postura puede ser resumida así:

In general, Hooke's attitude to Newton's theory was the following: he was willing to grant the experimental results reported by Newton, but was unable to accept the 'hypothesis' which the latter had proposed to explain them; he further expressed his belief that all of Newton's experiments, including the *experimentum crucis*, could be equally well explained by his own hypothesis about the nature of light and colours (Sabra, 1981, pp. 251-252).

Y, de este modo, su crítica a la teoría de Newton sobre la luz y los colores se reduce a profesar su acuerdo con los resultados de los experimentos que Newton realiza y a rechazar las hipótesis utilizadas por Newton para “salvar los fenómenos”:

But, tho' I wholly agree with him as to the truth of those [observaciones] he hath alledged, as having, by many hundreds of trials, found them so; yet as to his hypothesis of solving the phenomæna of colours thereby, I confess, I cannot see yet any undeniable argument to convince me of the certainty thereof (Hooke, 1672, pp. 10-11).

Para Hooke, los experimentos que Newton alega como pruebas de su “nueva teoría” podrían ser interpretados satisfactoriamente desde su propia hipótesis acerca de la luz, según la cual ésta consiste en pulsos de un medio

For all the experiments and observations I have hitherto made, nay, and even those very experiments, which he alledged, do seem to me to prove, that white is nothing but a pulse or motion, propagated through an homogeneous, uniform and transparent medium: and that colour is nothing but the disturbance of that light, by the communication of that pulse to other transparent mediums, that is, by the refraction thereof (*ibid.*, p. 11).

Hooke rechaza por lo tanto las consideraciones de Newton acerca de la naturaleza heterogénea de la luz, proposición que Newton considera probada por su *experimentum crucis* y que sin embargo a Hooke no le convence porque no ve la necesidad de que la diferente refrangibilidad de los rayos tras la refracción tenga que ser entendida como una propiedad original de los rayos individuales.

That the ray of light is as it were split or rarified by refraction, is most certain; [...].

But why there is a necessity, that all those motions, or whatever else it be that makes colours, should be originally in the simple rays of light, I do not yet understand the necessity (*ibid.*).

Para Hooke, la luz, antes de atravesar un medio refractante, consiste en un movimiento simple y uniforme que se propaga con igual velocidad. Cuando es refractada por un prisma, se generan múltiples movimientos o pulsos oblicuos y debido a ellos aparecen los colores. Seguro de que su teoría es mejor y más simple para dar cuenta de estos fenómenos ópticos, Hooke menosprecia la importancia que Newton atribuye al *experimentum crucis*, de tal modo que concluye su crítica a la “New Theory” reafirmando su propia hipótesis acerca de la naturaleza de la luz y negando que Newton haya realizado ninguna demostración “absoluta” de su nueva teoría:

[...] light is nothing but a simple and uniform motion, or pulse of a homogeneous and adopted (that is a transparent) medium, propagated from the luminous body in orbem, to all imaginable distances in a moment of time, and that that motion is first begun by some other kind of motion in the luminous body; [...]. I believe Mr. NEWTON will think it no difficult matter, by my hypothesis, to solve all the phænomena [...]. If Mr. NEWTON hath any argument, that he supposes as absolute demonstration of his theory, I should be very glad to be convinced by it (*ibid.*, pp. 14-15).

La crítica realizada por Hooke le parecerá injusta a Newton. Según éste, Hooke cree que toda la teoría newtoniana sobre la luz y los colores se basa en una sola



SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

hipótesis, la de la naturaleza corpuscular de la luz, de la cual dependen el resto de proposiciones afirmadas por él. Sin embargo, en opinión de Newton, la hipótesis corpuscular no es más que una hipótesis sin confirmar que efectivamente él ha sugerido como probable pero que no forma parte de las propiedades de la luz afirmadas en su doctrina positiva.

I shall now take a view of the Considerations on my Theories. And those consist in ascribing an Hypothesis, which, as to the principal parts, is not against me; in Granting the greatest part of my discourse if explicated by that Hypothesis; and in Denying some things, the truth of which would have appear'd by an experimental examination.

[...] Of these Particulars I shall discourse in order. And first of the Hypothesis, which is ascribed to me in these words: But grant his first supposition, that light is a body, and that as many colours or degrees as there may be, so many bodies there may be; all of which compounded together would make White, &c. This, it seems, is taken for my Hypothesis. 'Tis true, that from my Theory I argue the Corporeity of Light; but I do it without any absolute positiveness, as the word perhaps intimates; and make it at most but a very plausible consequence of the Doctrine, and not a fundamental Supposition, nor so much as any part of it (Newton, 1672 [respuesta a Hooke], p. 5086).



Por lo tanto, Newton niega que la naturaleza corpuscular de la luz sea, como sugería Hooke, la primera suposición de su teoría y trata de delimitar lo que son las proposiciones derivadas a partir de experimentos de las hipótesis. Así, haciendo constar que las propiedades de la luz que ha establecido en su nueva doctrina están probadas independientemente de cualquier hipótesis relativa a la naturaleza de la luz, Newton admite que esas propiedades, una vez que han sido demostradas, pueden ser explicadas tanto por su hipótesis corpuscular como por hipótesis mecánicas, entre las que puede hallarse la hipótesis de que la luz consiste en ciertos pulsos o movimientos del medio. Pero estas explicaciones estarían dentro de un ámbito especulativo, que no hay que mezclar con el de la ciencia positiva.

Had I intended any such Hypothesis, I should somewhere have explain'd it. But I knew, that the Properties, which I declar'd of Light, were in some measure capable of being explicated not only by that, but by many other Mechanical Hypotheses. And therefore I chose to decline them all, and to speak of Light in general terms, considering it abstractly, as something or other propagated every way in streight lines from luminous bodies, without determining, what that Thing is; whether a confused Mixture of difform qualities, or Modes of bodies, or of Bodies themselves, or of any Virtues, Powers, or Beings whatsoever (*ibid.*, pp. 5086-5087).



Newton se niega a que sus adversarios rechacen las propiedades que él ha considerado demostradas tras haber relatado su *experimentum crucis*: que la luz consiste en un agregado heterogéneo es algo que Newton no considera una hipótesis, aunque sus rivales no opinen lo mismo. La estrategia de Newton para establecer su teoría como probada y rechazar las hipótesis rivales consistía en argumentar mediante experimentos, de manera que las proposiciones deducidas de éstos fuesen aceptadas objetivamente, y evitar que cualquier hipótesis (del tipo que fuese) condicionase esa argumentación. La aceptación de que todo movimiento ha de ser explicado mecánicamente era para Newton un principio *a priori* que él no compartía y creía que esa hipótesis condicionaba la aceptación por parte de sus rivales de ciertas propiedades demostradas experimentalmente. En el caso de Hooke, éste aseguraba que esa hipótesis podía dar cuenta del fenómeno observado por Newton, ¿pero era eso consistente con su teoría mecanicista y modificacionista? Lo que Newton pretende demostrar en su *experimentum crucis* es la no modificabilidad de los rayos en una segunda refracción. Hooke acepta que los rayos sean separados por refracción en el primer prisma, eso también lo había observado él, pero lo que no acepta es que las propiedades de los rayos que provocan los colores (fueran éstas lo que fueren) tengan que ser propiedades originales a los rayos de luz. No obstante, si no son propiedades originales y son “creadas” en la refracción a través del prisma, debido a la oblicuidad que se produce en el movimiento de los rayos al cambiar de medio, no se explica por qué esos rayos no vuelven a cambiar su dirección al atravesar el segundo prisma, ya que los principios por los que ocurre lo primero serían aplicables a una segunda refracción. Además, esto parece contradictorio con la afirmación que realiza Hooke en su *Micrografía*, según la cual admite que una segunda refracción sí actuaría sobre los rayos que han sufrido una primera refracción: «Así pues, habiendo hallado que esta extraña propiedad [la oblicuidad de los pulsos] es un concomitante inseparable de un rayo refractado que no esté enderezado por una refracción contraria, examinaremos a continuación [...]» (Hooke, 1989, p. 250). Esto lo afirmaría en referencia al restablecimiento de la luz blanca tras haber sido separada en rayos de diferentes colores (Hooke utilizaba un frasco de laboratorio para separar los rayos y solo obtenía bien separados los extremos rojo y azul). No obstante, Hooke no afirma haber realizado ese experimento de recomposición de la luz blanca, con lo que parece estar conjeturando lo que ocurriría.

Newton y Hooke no llegarían a ponerse de acuerdo. La rivalidad entre ambos quizás excedía lo puramente teórico y, como sugiere Shapiro (2008), la lucha por una posición de autoridad estaba como telón de fondo: «Newton was clearly affronted by Hooke's sharp criticism, but there is more to the story than that. Hooke was then England's leading authority on optics» (p. 419). Newton se proponía llevar a cabo el programa que Hooke había establecido como necesario para ofrecer un tratamiento completo de

SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

la óptica moderna, un programa que diese cuenta de la luz y los colores no solo mediante la explicación de la refracción, sino también de otros fenómenos como por ejemplo los colores de las láminas finas y la difracción. Así, ante las críticas que recibe de Hooke, es consciente de que no basta con discutir en términos apriorísticos advirtiendo del peligro de mezclar las hipótesis con las proposiciones demostradas, sino que ve necesario reunir la mayor evidencia empírica posible para convencer a sus rivales de su teoría.

On 11 May 1676 he wrote Oldenburg that he still hoped to be able to get some time in the summer to write this new treatise "about ye colours of ye Prism" [...] Newton was evidently now planning a work very much like the *Opticks* [...]. The debate over the "New Theory" with Hooke, Huygens, and others had compelled him to recognize that significant parts of this theory of color had to be recast and more rigorously demonstrated (Shapiro, 2008, p. 425).

8. LA CRÍTICA DE PARDIES

Al igual que Hooke, Pardies también entiende que la afirmación de la heterogeneidad de la luz por parte de Newton va ligada a una concepción corpuscular de la naturaleza de la luz. Así, en su primera carta a Oldenburg criticando la nueva teoría de Newton, Pardies se refiere a esa teoría como una ingeniosa hipótesis:

I have read Mr. Newton's very ingenious hypothesis of light and colours [...].

It seems very extraordinary that the learned author should make light to consist of an almost infinite number of rays, endued with a natural disposition of retaining and exhibiting their own proper colours, and that are disposed in a certain peculiar way to be refracted, some in a greater, and others in a less degree (Pardies, 1958 [primera carta], pp. 86-87).

Newton contestará ante esa calificación de su teoría como hipótesis, mostrando su desacuerdo. Para Newton, ese término tenía una acepción negativa y no era aplicable a una teoría probada por los experimentos, como era el caso de su nueva teoría sobre la luz y los colores. Sin embargo, Pardies no utilizaba ese término pensando en esa acepción específica: lo que parecía significar era que reconocía que la heterogeneidad de la luz blanca era el núcleo de la teoría de Newton y que, por supuesto, no creía que su defensa estuviese justificada. La respuesta de Newton pasaba por reafirmar el carácter probado y verdadero de las propiedades de la luz establecidas por él.

I do not take it amiss that the Rev. Father calls my theory an hypothesis, inasmuch as he was not acquainted with it. But my design was quite different, for it seems to contain only certain properties of light, which, now discovered, I think easy to be proved, and which if I had not considered them as true, I would rather have them rejected as vain and empty speculation, than acknowledged even as an hypothesis (Newton, 1958 [respuesta a la primera carta de Pardies], p. 92).

www.solofici.org

Pero la crítica de Pardies no se restringía a tachar la teoría de Newton como hipótesis. Pardies percibía que la teoría de Newton socavaba los fundamentos de la dióptrica al desafiar el postulado según el cual las leyes de la refracción serían las mismas para todos los rayos de luz. Pardies creía que no era necesario postular una diferencia en la refrangibilidad de los rayos, ya que él pensaba que la figura oblonga formada en la pared podía ser explicada por las reglas de la refracción, asumiendo que los rayos son igualmente refrangibles.

But it appears to me that, according to the common and received laws of dioptrics, the figure ought to be, not round but oblong. For since the rays proceeding from the opposite parts of the sun's disk, are variously inclined in their passage to the prism, they ought also to be variously refracted; that since the inclination of some rays is at least 30' more than that of others, their refraction must also be greater (Pardies, 1958 [primera carta], p. 87).

Por lo tanto, para Pardies, asumiendo las leyes tradicionales de la refracción según las cuales todos los rayos tienen la misma refrangibilidad, sería la diferencia en el ángulo de incidencia la que daría lugar a la figura oblonga: «[...] when the same rays on the superior and inferior parts, are refracted, some more, some less, as being unequally inclined, they must needs diverge more, and consequently be extended in an oblong figure» (*ibid.*, p. 88). No habría necesidad, por lo tanto, de desafiar ese postulado, según el cual el *experimentum crucis* también puede ser explicado: «What the author calls the Experimentum Crucis, seems also to agree with the commonly received laws of refraction» (*ibid.*).

It is not then to be wondered at, if these rays falling severally on a second prism, and having a very small hole in it, be unequally refracted, since their inclination is unequal. Nor does it alter the case, that those rays are raised or depressed by the rotation of the first prism, the second remaining immoveable, [...] that it may successively receive the coloured rays of the whole image, and transmit them through its proper hole (*ibid.*, p. 89).



SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

De este modo, Pardies realiza una serie de mediciones para demostrar esto pero pasaba por alto una advertencia de Newton a la hora de disponer el prisma en la pared, indicada antes de describir el *experimentum crucis*: éste ha de ser dispuesto de tal modo que se minimice la diferencia entre los rayos de sol que provienen de las diferentes partes de éste (y por tanto, cuyos ángulos de incidencia varían). Así explicaba Newton cuál era el error de Pardies:

But the Rev. Father is under a mistake. For he has made the refractions by the different parts of the prism to be as unequal as possible, whereas in the experiments, and in the calculation from them, I employed equal refractions (Newton, 1958 [respuesta a la primera carta de Pardies], p. 90).

A la hora de describir su *experimentum crucis*, Newton daba por supuesta esta aclaración previa y no explicitaba de qué modo exacto había que disponer el prisma, con lo cual, en cierto modo es comprensible que, debido a la escueta descripción de su experimento, sus lectores no entendiesen perfectamente cómo había que disponer los instrumentos. Newton explicita que si los ángulos de incidencia de todos los rayos son iguales, el problema de la figura oblonga sigue manifestándose, con lo cual el *experimentum crucis* sirve para probar la diferente refrangibilidad de esos rayos.

What the Rev. Father contends is not inconsistent with what I called the Experimentum Crucis, viz. that the unequal refractions of rays endued with different colours, were produced by unequal incidences: for transmitting rays through two very small immoveable holes, and at a distance from each other, the incidences, as I made the experiment, were always equal, and yet the refractions were manifestly unequal (*ibid.*, p. 91).

Pero en esa primera carta de Pardies, éste no solo criticaba la heterogeneidad de la luz blanca, sino que también atacaba la concepción de los colores de Newton según la cual la mezcla de todos los que habían sido descompuestos a través del prisma daría lugar de nuevo a la luz blanca.

The author's notion of colours indeed follows very well from the preceding hypothesis; yet it is not without its difficulties. For when he says, that all the rays being promiscuously blended together, yield no colour, but rather a whiteness, this does not seem conformable to all the phaenomena. [...] the red, yellow, blue, purple, and all the others, when mixed together, produce, not a white, but an obscure sated colour. So also ordinary light should appear of the same colour, being a like aggregate of all colours (Pardies, 1958 [primera carta], p. 89).

Para Pardies, en la mezcla de rayos de colores ocurriría algo similar a lo que ocurre en la mezcla de tintas de colores, que no producen un blanco sino un color oscuro. Para Newton, la diferencia entre los colores radica en la diferente cantidad de blanco de la que está compuesto cada uno.

Against the theory of colours it is objected, that powders of divers colours mixed together, do not yield a withe, but an obscure and dusky colour. But to me, white, black, and all the intermediate dusky colours, which can be compounded of mixtures of white and black, do not differ as to their species, but only as to their quantity of light. (Newton, 1958 [respuesta a la primera carta de Pardies], p. 91).

En una segunda carta de Pardies, éste admite las explicaciones de Newton sobre el ángulo de incidencia de los rayos y sobre los colores. Pero sigue cuestionando que la diferente refrangibilidad de los rayos constituya la única manera posible de dar cuenta de la figura oblonga, y a ese respecto sugiere que otras posibles explicaciones también podrían dar cuenta de los fenómenos, como la de Grimaldi o la de Hooke:

[...] the greater length of that image may be otherwise accounted for, than by the different refrangibility of the rays. For according to that hypothesis, which is explained at large by Grimaldi, and in which it is supposed that light is a certain substance very rapidly moved, there may take place some diffusion of the rays of light after their passage and decussation in the hole. Also on that other hypothesis, in which light is made to proceed by certain undulations of a subtile matter, as explained by Mr. Hook, colours may be explained by a certain diffusion and expansion of the undulations, made on the sides of the rays beyond the hole by the influence and continuation of the subtile matter. (Pardies, 1958 [segunda carta], p. 104).

La respuesta de Newton pasará por reconocer primero la posibilidad de la que habla Pardies, esto es, que otras teorías podrían explicar el fenómeno de la figura oblonga, para después asegurar que precisamente lo que ha tratado de hacer es refutar dichas teorías. Así, Newton (1958 [respuesta a segunda carta de Pardies]) dice que a las hipótesis de la difusión que cita Pardies, la de Grimaldi, la de Hooke, también podría añadirse la de Descartes: «in which a similar diffusion of *conatus*, or pression of the globules, may be conceived» (p. 106). Y continúa admitiendo que esa difusión o expansión puede ser explicada por diferentes hipótesis «in which light is supposed to be a power, action, quality, or certain substance emitted every way from luminous bodies» (*ibid.*). Sin embargo, a continuación, Newton recalca la diferencia



SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

entre las propiedades demostradas por experimentos y las hipótesis, no demostradas, para volver a insistir en que su teoría forma parte de las primeras y no de las segundas. Una vez aceptadas las propiedades de la luz que han sido probadas experimentalmente, si se quiere, se puede conjeturar acerca de las hipotéticas causas de esas propiedades, pero al margen de lo probado científicamente.

In answer to this, it is to be observed that the doctrine which I explained concerning refraction and colours, consists only in certain properties of light, without regarding any hypotheses, by which those properties might be explained. For the best and safest method of philosophizing seems to be, first to inquire diligently into the properties of things, and establishing those properties by experiments and then to proceed more slowly to hypotheses for the explanation of them. For hypotheses should be subservient only in explaining the properties of things, but not assumed in determining them; unless so far as they may furnish experiments. For if the possibility of hypotheses is to be the test of the truth and reality of things, I see not how certainty can be obtained in any science; since numerous hypotheses may be devised, which shall seem to overcome new difficulties (*ibid.*).

Newton explica que la argumentación que lleva a cabo en su artículo "New Theory of Light and Colors" consta de una serie de premisas (que entroncan con los conocimientos de trasfondo aceptados por la comunidad científica), y de una serie de experimentos. Newton explicita a Pardies las premisas por las cuales emplea una determinada definición de lo que es la luz y lo que son los rayos (sus últimas partes, independientes), sin que estas definiciones conlleven ninguna proposición relativa a la naturaleza física de la luz.

By light therefore I understand, any being or power of a being, (whether a substance or any power, action, or quality of it, which proceeding directly from a lucid body, is apt to excite vision. And by the rays of light I understand its least or indefinitely small parts, which are independent of each other; such as are all those rays which lucid bodies emit in right lines, either successively or all together (*ibid.*).

Newton considera que esas premisas sobre lo que son la luz y los rayos no le obligan a asumir (ya que no implican) ninguna hipótesis particular sobre la naturaleza física de la luz (si ésta es corpuscular u otra cosa). Una vez asumidas esas definiciones, la única explicación alternativa a la que él propone para explicar el problema del espectro oblongo, reconoce que es la misma que sugiere Pardies, a saber, que la figura oblonga se deba a una difusión de los rayos que tuviese lugar "más allá" del agujero. Pero

precisamente porque era consciente de esta posibilidad, en su artículo de 1672, antes de anunciar su nueva teoría sobre la luz y los colores, describe una serie de experimentos que van destinados a mostrar que la diferente forma en que son refractados los rayos por el prisma no se debe a ningún tipo de irregularidades que pudiesen ser la causa de tal difusión.

www.solofici.org



This being premised, the whole force of the objection will lie in this, that colours may be lengthened out by some certain diffusion of light beyond the hole, which does not arise from the unequal refraction of the different rays, or of the independent parts of light. And that the image is no otherwise lengthened, was shown in my letter in Numb. 80 of the Transactions; and to confirm the whole in the strictest manner, I added that experiment now known by the name *Experimentum Crucis* (*ibid.*, pp. 106-107).



Por lo tanto, la causa de la figura oblonga no sería la difusión de la que habla Pardies, sino la diferente refrangibilidad de los rayos, como confirma el *experimentum crucis*. Así, tras unas aclaraciones sobre cómo disponer bien dicho experimento, Newton concluye:



This shows that the diffusion, above-mentioned, does not arise from the influence or continuity of the undulating matter, or matter put into a rapid motion, or any such like causes, but from a certain law of refractions for every species of rays (*ibid.*, p. 108).



Orientando su "New Theory" a refutar experimentalmente las explicaciones modificacionistas sobre la causa de la figura del espectro, la fuerza probatoria a favor de la teoría de Newton se concentraba en un único *experimentum crucis* que quizás no estaba explicado con el detalle suficiente que se requeriría para poder realizar una réplica exacta de él sin que se interpusiesen ciertos malentendidos técnicos. Así, debido a la falta de una descripción detallada de la disposición y de las características de los instrumentos empleados se produjeron varios intentos de repetición del mismo que resultaron fallidos. En esta línea, S. Schaffer (1989) llama la atención sobre el uso del prisma como pieza angular de la demostración de la nueva teoría newtoniana sobre la luz y los colores. Según este autor, Newton obvia ciertos detalles sobre este instrumento que son necesarios, y hasta pueden ser decisivos, para poder replicar con éxito sus experimentos. Pero como indica Shapiro (1996), las dificultades para replicar el experimento crucial de Newton rebasarían la mera falta de detalles acerca del instrumento físico clave, el prisma, siendo necesario entender el cambio que Newton realiza en la interpretación de lo que ocurre en el experimento respecto a lo que se venía estableciendo desde la tradición óptica generalmente aceptada. Ya no se trata de explicar cómo el prisma genera los colores: se trata de entender que los colores no son generados por el prisma sino simplemente separados por él.

SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

Newton had to teach the scientific community a new way of thinking about what a prism does and change it from a generator of colors that merely deflects the image to a separator of colors that operates by unequally refracting and dispersing the rays. This was a significant conceptual stumbling block. The problem for Newton's contemporaries was not with the simple physical apparatus, as Schaffer holds, but with his exposition (Shapiro, 1996, p.72).

Por lo tanto, vemos que la controversia entre Newton y sus críticos gira en torno a la aceptación de la interpretación teórica que Newton realiza del *experimentum crucis*. Newton considera justificada esa interpretación debido al método científico según el cual las propiedades de las cosas han de deducirse de los experimentos. Pero sus adversarios no lo ven tan fácil, pues esa nueva forma de interpretar se opone a la tradición establecida, de la que ellos son continuadores. Aceptar, como pretende Newton, que los rayos de luz no son igualmente refrangibles implica aceptar que la luz no es homogénea, socavando los cimientos de las teorías modificacionistas tal como se entendían en ese momento.



9. LAS RESERVAS DE HUYGENS ANTE LA TEORÍA DE NEWTON

En 1672, Oldenburg envía el artículo de Newton a Huygens para que éste proporcione su opinión al respecto. En una primera respuesta de Huygens, éste ignora completamente el núcleo de la teoría expuesta por Newton, por el que esa teoría resultaba novedosa, y simplemente comenta que le agrada saber que Newton ha percibido, al igual que él ha hecho ya, que la refracción de ciertos cristales “debido a la inclinación de sus superficies” afecta a los telescopios. Respecto a su “nueva teoría sobre los colores” dice que le parece ingeniosa pero que será necesario ver si concuerda con los experimentos. El desdén de Huygens por la nueva propuesta de Newton es notable. Pero si bien es cierto que Newton hacía referencia en su artículo al fenómeno de la aberración cromática, él no creía que ese problema se debiese a la forma de los cristales, sino a la constitución misma de la luz, de acuerdo con la cual ésta consiste en rayos de refrangibilidad variable:

I saw, that the perfection of Telescopes was hitherto limited, not so much for want of glasses truly figured according to the prescriptions of Optick Authors, (which all men have hitherto imagined,) as because that Light it self is a Heterogeneous mixture of differently refrangible Rays. So that, were a glass so exactly figured, as to collect any one sort of rays into one point, it could not collect those also into the same point, which having the same Incidence upon the same Medium are apt to suffer a different refraction (Newton, 1671/2 [“New Theory”] p. 3079).



Tras una segunda petición de Oldenburg para que Huygens comentase la nueva teoría de Newton, éste pasa de ignorarla a calificarla de hipótesis probable, pero sin mencionar el núcleo de la nueva teoría: la diferente refrangibilidad de los rayos. Más tarde, Oldenburg envía a Huygens una copia de la correspondencia mantenida entre Newton y Pardies y de un nuevo artículo de Newton explicando su teoría (“A Serie's of Quere's”⁷), con lo cual es de suponer que Huygens sí estaba al tanto de en qué consistía el núcleo de esa nueva teoría y de la polémica suscitada. Sin embargo, sus reservas respecto a la novedosa teoría continúan y en referencia a ese nuevo artículo de Newton en el que éste manifiesta que los experimentos confirman su teoría “positivamente y directamente”, Huygens dirá que Newton debería presentar esa teoría como una hipótesis muy plausible. La explicación que Newton hace de los colores (sigue sin mencionar la diferente refrangibilidad), no le parece satisfactoria desde el punto de vista de la filosofía mecánica. Como dice Sabra (1981): «As a mechanist and a Cartesian, he was not satisfied with Newton's doctrine as long as it did not make it clear in what sense colours are original properties of the rays» (p. 270). Huygens cree que si admitiésemos que cada rayo tiene un color desde su origen, todavía quedaría por explicar mecánicamente en qué consiste esa diversidad de colores.

En una subsiguiente carta de Huygens a Oldenburg, el primero por fin se atreve a mencionar la diferente refrangibilidad de los rayos pero insiste en la ausencia de una explicación mecánica para los colores. Huygens redacta esta carta tras haber leído la respuesta de Newton a la crítica de Hooke. Así, siguiendo a Hooke, indica que la objeción más grande a la nueva teoría de Newton es la dificultad de dar cuenta de la luz blanca por medio de tantos colores primarios como considera Newton que la componen. Huygens cree que esa explicación es mucho más sencilla desde su propia teoría, según la cual solo se admitirían dos colores primarios, el amarillo y el azul:

I Have seen, how Mr. Newton endeavours to maintain his new Theory concerning Colours. Methinks, that the most important Objection, which is made against him by way of Quære, is that, Whether there be more than two sorts of Colours. For my part, I believe, that an Hypothesis, that should explain mechanically and by the nature of motion the Colours Yellow and Blew, would be sufficient for all the rest, in regard that those others, being only more deeply charged (as appears by the Prismes of Mr. Hook,) do produce the dark or deep-Red and Blew; and that of these four all the other

7. El título completo del artículo reza así: “A Serie's of Quere's propounded by Mr. Isaac Newton, to be determin'd by Experiments, positively and directly concluding his new Theory of Light and Colours; and here recommended to the Industry of the Lovers of Experimental Philosophy, as they were generously imparted to the Publisher in a Letter of the said Mr. Newtons of July 8. 1672”.

SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

colors may be compounded. Neither do I see, why Mr. Newton doth not content himself with the two Colors, Yellow and Blew; for it will be much more easy to find an Hypothesis by Motion, that may explicate these two differences, than for so many diversities as there are of other Colors. And till he hath found this Hypothesis, he hath not taught us, what it is wherein consists the nature and difference of Colours, but only this accident (which certainly is very considerable,) of their different Refrangibility (Huygens, 1673, p. 6086).

Newton entiende, por la respuesta de Huygens, que esos dos colores, el amarillo y el azul, tendrían que bastar para dar cuenta de la composición de la luz blanca, lo cual cree que no hace justicia a lo observado por él al descomponer la luz del sol mediante el prisma, apareciendo todos los colores del espectro. Así, tras varios malentendidos con Huygens respecto a este asunto, Newton escribe una carta aclarando punto por punto su teoría, y por lo que respecta a la principal objeción de Huygens respecto a la explicación de los colores, reafirmará su concepción de la luz blanca:

Whiteness in all respects like that of the Sun's immediate light and of all the usual objects of our senses cannot be compounded of two simple colors alone. For such a composition must be made by rays that have only two degrees of refrangibility [...].

The Sun's light is an aggregate of an indefinite variety of homogeneous colors [...]. And hence it is, that I call homogeneous colors also primitive or original. And thus much concerning Colors. (Newton, 1673 [respuesta a Huygens], pp. 6090-6091).



Newton intenta zanjar la crítica realizada por Huygens, que parece considerar en cierto modo vaga y poco concreta. Pero, su respuesta no acaba de satisfacer a Huygens, para quien Newton no logra explicar mecánicamente mediante su nueva teoría en qué consiste la naturaleza de los colores y su diversidad. Al definir los colores como cualidades de los rayos y no aventurarse a asegurar en qué consistía la naturaleza de éstos, Newton se mantenía precavido respecto a lo que él consideraba hipótesis. En relación a esto, ya había justificado así su postura ante Hooke:



But I knew, that the Properties, which I declar'd of Light, were in some measure capable of being explicated not only by that, but by many other Mechanical Hypotheses. And therefore I chose to decline them all, and to speak of Light in general terms, considering it abstractly, as something or other propagated every way in streight lines from luminous bodies, without determining, what that Thing is; whether a confused Mixture of difform qualities, or Modes of bodies, or of Bodies themselves, or of any Virtues, Powers, or Beings whatsoever. And for the same reason I chose to speak of Colours according to the information of our Senses, as if they were Qualities of Light without us. Whereas by that Hypothesis I must have considered them rather as Modes of Sensation, excited in the mind by various motions, figures, or sizes of the corpuscles of Light, making various Mechanical impressions on the Organ of Sense (Newton, 1672 [respuesta a Hooke], pp. 5086-5087).

Newton dice que ha decidido hablar de los colores “como si fuesen cualidades de la luz sin nosotros”, ya que considera que son propiedades objetivas de la luz al margen de nuestros sentidos. En la *Óptica*, aclarará en qué sentido podemos entender esas propiedades objetivas.

Si en algún momento hablo de luces o rayos de colores o digo que están dotados de colores, ha de entenderse que no estoy hablando filosóficamente o con propiedad, sino groseramente y según esos conceptos que las personas ordinarias habrían de tramar frente a todos estos experimentos, pues, propiamente hablando, los rayos no tienen colores. En ellos, no existe más que una capacidad o *disposición* para despertar este o ese color [cursivas añadidas] (Newton, 1672, pp. 112-113).

Pero aún hablando en términos filosóficos y entendiendo los colores como esas “disposiciones” de los rayos, la demanda de Huygens de una explicación mecánica de los rayos seguiría sin ser satisfecha, ya que todavía es posible preguntar en qué consisten esas “disposiciones”, en términos físicos.

Newton, however, always refrained from offering a direct and definite answer to Huygens' question. Just as he wished to speak of colours in 1672 'according to the information of our senses', he also preferred to speak in the *Opticks* (1704) 'not philosophically and properly' but as the 'vulgar' would speak. He knew that speaking philosophically would inevitably be in terms of a particular view about the nature of light. But as this would explicitly tie up his theory with a contestable hypothesis, he preferred to expound the allegedly experimentally proved properties of light in dispositional terms (Sabra, 1981, pp. 293-294).

SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

Así, Newton diferenciaba entre esa consideración objetiva de las propiedades de los rayos, que se deducen de los fenómenos y son independientes de cualquier hipótesis sobre la naturaleza física de la luz, y la descripción de cómo los rayos producen las sensaciones en nuestra visión, lo cual sí conllevaría entenderlos en términos de corpúsculos. Es por ello que Newton reserva este tema para las “Cuestiones” de la *Óptica*, conjeturando acerca de cómo los rayos excitarían los nervios ópticos para producir las diferentes sensaciones de los colores.

¿Acaso los rayos de luz no excitan vibraciones en la *Túnica Retina* cuando inciden sobre el fondo del ojo? Tales vibraciones, al propagarse hasta el cerebro por las fibras sólidas de los nervios ópticos, producen el sentido de la visión.



¿Acaso los diversos tipos de rayos no producen vibraciones de diversas magnitudes y, de acuerdo con ellas, excitan sensaciones de diversos colores [...]? (Newton, 1977, pp. 299-300).



En conclusión, hemos visto cómo los rivales de Newton usaban el término “hipótesis” como arma, para calificar como tal a la nueva teoría de Newton y admitir así su posibilidad pero no su verdad. En relación a esto, Sabra (1981) cree que tanto Hooke, como Pardies y Huygens siguen un mismo patrón:

It will have been noticed that the objections of Hooke, Pardies and Huygens exhibited somewhat the same pattern. They all understood Newton to be propounding a corpuscular view of light from which, they all granted, the rest of Newton's doctrine of colours could be deduced. They referred to Newton's theory as a hypothesis, thereby implying that the experiments could be interpreted in other ways. Being themselves inclined towards a view according to which light consisted in the transmission of motion through a medium rather than the transport of body, they either proposed alternative hypotheses (Hooke and Pardies) or expressed reservations regarding the Newtonian theory (Huygens) (p. 273).

Newton cree que las propiedades de la luz deberían ser investigadas independientemente de cualquier hipótesis sobre su naturaleza y considera que su doctrina está firmemente fundada en los experimentos, por lo que no estima necesaria (para la aceptación de dicha teoría) ninguna ulterior explicación mediante una hipótesis sobre la naturaleza de la luz. Una cosa son las conjeturas que cada uno haga y otra cosa tratar de imponer una hipótesis concreta en contra de lo que muestran los experimentos. «It was in connection with this claim that Newton made his first pronouncements on the proper method of scientific procedure. According to him, this method consisted in deducing the properties of thing from experiments; not in arguing from hypotheses» (*Íbid.*, p. 274).

A partir de ese momento, su objetivo no será otro que mostrar a sus oponentes que su teoría se sigue fielmente de lo que muestran los experimentos y no ha de ser rechazada por ningún tipo de hipótesis.

10. EL INTENTO DE REPETICIÓN FALLIDO DE LUCAS.



Aparte de las controversias con Hooke, Pardies y Huygens, el *experimentum crucis* de Newton también provocó una discusión por correspondencia entre Newton y un grupo de jesuitas ingleses afincados en Lieja: Linus, Gascoines y Lucas. Shapiro (1996) no presta mucha atención a estas discusiones debido a la poca relevancia que tenía este grupo dentro de la comunidad científica. Según él, se les habría tachado de incompetentes al no haber sido capaces de replicar los experimentos de Newton, empezando por el más básico según el cual se obtenía la figura oblonga del espectro en la pared tras descomponer los colores a través del prisma. En las primeras cartas que intercambiaron con Newton sobre este asunto, negaban que eso ocurriese como describía Newton, a lo cual Newton respondía que tenían que estar realizando mal el experimento (o no realizándolo):

For this is to be decided not by discourse, but new tryall of ye experiment. What it yt imposes upon Mr Line I cannot imagin, but I suspect he has not tryed ye expt since he acquainted himself wth my Theory, but depends upon his old notions taken up before he had any hint given to observe ye figure of ye coloured image. I shall desire therefore, before he returns any answer, to try it once more for his satisfaction & that according to this manner. (Newton, 1959 [“143 Newton to Oldenburg”], pp. 356-357).

Después de explicar el procedimiento correcto para que Linus pudiese obtener la figura oblonga en la pared, Newton le anima a que realice el *experimentum crucis*, ya que, al igual que con los demás experimentos, había negado sus resultados.

When Mr Line has tryed this I could wisht he would proceed a little further to try yt wch I called ye *Experimentum Crucis*, seing (if I misremember not) he denies that as well as the other (wch by his denying them I know he has not done yet as they should be tryed) I presume he will rest satisfied (*ibid.*, p. 357).

A pesar de que Newton les explicó detalladamente cómo disponer los instrumentos en el experimento básico para obtener la figura oblonga en la pared, los jesuitas de Lieja aseguraban que Linus sí había repetido los experimentos de Newton obteniendo resultados diferentes, por lo que a no ser que hubiese alguna diferencia técnica a la hora de disponer los instrumentos, los resultados de Newton no se sostenían.

SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

But what Mr Newton fancy's, as if Mr Line depended rather upon old experiments than any new trial made since the beginning of this debate, is a thing much contrary to Mr Line's known humour and practice, and what many here are witness of. [...] So, that for point of experience, Mr Newton cannot be more confident on his side, than we are here on the other, who are fully persuaded that unless the diversity of placing the Prisme, or the bigness of the hole, or some other such circumstance, be cause of the difference betwixt them, Mr Newton's experiment will hardly stand. For as Mr Line was always at home, and in his chamber, and ordinarily kept his Prisme just ready before the hole, so we think it probable he hath tried his experiment thrice for Mr Newton's once, and that in a clearer and more uncloudy sky than ordinarily England doth allow (Gascoines, 1959 ["148 Gascoines to Oldenburg"], p. 394).

Cuando el experimento básico de la descomposición de los colores mediante el prisma para obtener el espectro oblongo en la pared fue realizado en la Royal Society, coincidiendo el resultado con el que Newton había reportado, esto fue comunicado al grupo de Lieja, debido a que Linus negaba el resultado del experimento.

The experiment of Mr. Newton, which had been contested by Mr. Linl's and his fellows at Liege, was tried before the Society, according to Mr. Newton's directions, and succeeded, as he all along had asserted it would do: and it was ordered, that Mr. Oldenburg should signify this success to those of Liege, who had formerly certified, that if the experiment were made before the Society, and succeeded according to Mr. Newton's assertions, they would acquiesce, as appears by Mr. Gascoigne's letter to Mr. Oldenburg of 15th December, 1675 (Birch, 1757, p. 313).

Tras esto, el sucesor de Linus, Lucas, respondería aceptando el resultado de la Royal Society pero interponiendo otra serie de objeciones a la teoría de Newton, en relación con su *experimentum crucis*.

Just upon the close of the adjoynd letter, I received from Mr Gascoine, yours of May the 4th; wherein you are pleased to favour us with an exact account of the famous experiment of the coloured spectrum, lately exhibited before the Royal Society. I was much rejoiced to see the tryalls of that Illustrious Company, agree soe exactly with ours here, thò in somewhat ours disagree from Mr Newton (Lucas, 1959 ["161 Lucas to Oldenburg"], p. 12).

Pero el grupo de Lieja seguía sin aceptar el resultado derivado por Newton del *experimentum crucis*. Lucas escribe nuevamente a Oldenburg negando que dicho experimento implique una diferente refrangibilidad para los rayos de colores diferentes.



[Newton] remitts me to the *experimentum Crucis* as a demonstrative prooffe of his new Assertion, and consequently as a sufficient solution to all my experimentall objections. [...] I cannot esteem it a demonstrative proof of this new Theory without the assistance of further experiments. This, I conceive may be made appeare by drawing the proof of this new Theory into a close Syllogisticall methode after this or the like tenour. *Rays differently coloured suffer an unequal refraction at an equall incidence, as appears from the experimentum Crucis: arqui this unequal refraction necessarily implyes an unequal refrangibility in rays differently coloured: ergo rays differently coloured have an unequal refrangibility at an equall incidence.* Now as this Syllogisme [...] the maine difficulty lyes in the Minor, viz *that this unequal refraction necessarily implyes an unequal refrangibility in rays differently coloured.* Whichs Minor is neither evident in it self [...] nor evidenced from the *experimentum Crucis*, this experiment at most proving an unequal refraction at an unequal refrangibility, *intrinsicall to rays differently coloured*, rather than from some extrinsecall cause (Lucas, 1959 ["185 Lucas to Oldenburg"], p. 104).



www.solofici.org

Tras negar que esa conclusión se siga del *experimentum crucis*, Lucas se dispone a repetir el experimento, hallando un resultado diferente al establecido por Newton, y demostrando así que no es posible afirmar una diferente refrangibilidad intrínseca a los rayos y que por lo tanto ésta no es la verdadera causa de la figura oblonga del espectro coloreado. Después de hacer pasar los rayos violetas a través del segundo prisma, encuentra una refracción mayor que la que ocurría con los rayos rojos al pasar por ese mismo prisma. Además, comprueba que esos rayos violetas van acompañados de una considerable cantidad de rojo, al refractarse tras el segundo prisma. Tras estas observaciones, concluye que el *experimentum crucis* muestra que no hay una correspondencia "uno a uno" entre colores y grados de refrangibilidad. Lucas describe así su réplica del *experimentum crucis*:



SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

In order then to this experiment I placed the first prisme as in my former letter: the 2d, distant from the window about 12 foot: the distance of the paper (wheron the image was cast) from the second prisme about 3 foot: the two boards at the prismes were placed exactly according to Mr Newtons directions. Now the result of many trialls was as follows. Upon casting the violet rayes on ye second prisme, I constantly indeed found a considerably greater refraction than when the red ones were cast thereon; yet withall (notwithstanding all care in excludeing extraneous light) I as constantly found thes violet rayes accompanied by a considerable quantity of red ones, upon ye paper behind the second prisme. This being so, I reason thus: Thes red rayes, notwithstanding the first refraction made in ye prisme at ye window, remained still interlaced with the violet ones; for, otherwise how should they constantly appeare behind the second prisme, upon casting the violet rayes thereon (*ibid.*, pp. 105-106).



Pero si el resultado reportado por Lucas era cierto, si efectivamente aparecía cierta cantidad de rayos rojos junto a los violetas tras la segunda refracción (la cual recordemos que debía actuar exclusivamente sobre una franja de rayos igualmente refrangibles), de esto se derivaría que el color es indiferente respecto al grado de refrangibilidad. Por lo tanto, no se seguiría ni la correspondencia unívoca entre color y grado de refrangibilidad, ni la inmutabilidad del color para cada rayo.

Whence it seems cleare 1° that the red rays suffered the same refraction in the first prisme at the window, as the violet ones did. 2° that the red rayes are of their owne nature wholly indifferent to the refractions of the red and violet. i. e. to the most opposit degrees of refractions in Mr Newtons Theory (*ibid.*, p. 106).

Newton niega que las observaciones de Lucas tras repetir el experimento crucial constituyan una prueba en contra de la inmutabilidad del color y alega que el experimento no ha sido replicado como es debido. En su carta a Lucas, tras citar el pasaje de la "New Theory" en necesidad de una perfecta separación de los colores, Newton (1959 ["220 Newton to Lucas"]) critica a Lucas por no haber logrado realizar esa perfecta separación entre los colores primarios, pero sin embargo no revela cómo se lograría ésta. «But since you know not how to make these separations, I perceive you never yet saw any of y(e) Primary colours about w(ch) you have written so much, nor know what they meane" (p. 258). Newton cree que Lucas confunde la diferente refrangibilidad de los rayos (unos rayos son más refrangibles que otros) con la correspondencia entre refrangibilidad y color, algo que no probaría el *experimentum crucis* en sí mismo sino que sería una cuestión posterior.

Having thus run over your letters let me add yt all your Objections run upon two general mistakes. You mistake both ye question of different refrangibility & ye notion of primary colours. [The Question of different refrangibility which I bring] ye *Experimentum Crucis* to decide, is not, as I sayd wither rays differently coloured are differently refrangible, but only whether some rays be more refrangible yn others. What ye colours of ye rays differently refrangible are, or whither they have any appropriate ones belongs to an after enquiry (*ibid.*, pp. 257-258).



Newton admite que tras su descripción del *experimentum crucis* en el artículo de 1672 afirma, como parte de su doctrina, que existe una correspondencia entre índice de refrangibilidad y color, y que tanto el color como el grado de refrangibilidad serían inmutables y propias de cada rayo.

To the same degree of Refrangibility ever belongs the same colour, and to the same colour ever belongs the same degree of Refrangibility. [...].

The species of colour, and degree of Refrangibility proper to any particular sort of Rays, is not mutable by Refraction, nor by Reflection from natural bodies, nor by any other cause, that I could yet observe (Newton, 1671/2 ["New Theory"], p. 3081).

Pero esa inmutabilidad del color tendría que ser demostrada por ulteriores experimentos, como asegura en su carta a Lucas, de tal modo que el *experimentum crucis* solo probaría la diferente refrangibilidad de los rayos.

www.solofici.org

You may see in my first Letter about Light (*Phil. Transact.* Num. 80) I make no mention of colours while I am prouving different refrangibility by ye *Experimentum Crucis*, but after yt begin to tel you yt ye rays wch differ in refrangibility differ also in colour, reserving this to be proved by other Experiments [...]. But you confound alltogether as if I brought ye *Experimentum Crucis* to prove all (Newton, 1959 ["220 Newton to Lucas"], p. 258).

Newton desecha las objeciones de Lucas según las cuales el resultado del *experimentum crucis* no afirmaría la correspondencia entre refrangibilidad y color, pues admite que ése no es el propósito de tal experimento. Por lo tanto, la única objeción que puede interponer Lucas sería la de que dicho experimento no tiene el mismo resultado que el reportado por Newton. Ante esta situación, uno de los dos tiene que estar realizando mal el experimento, y Newton cree que es Lucas.

SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

If you consider this you wil see yt while you have laboured to oppose different refrangibility, there's not one of your objections wch concerns it except ye third where ye Experiment succeeds otherwise than you have reported it. The rest are only against analogy between refrangibility & colour, for if I would say yt rays differently refrangible have no appropriate colours, all your objections would cease (*ibid.*).

Respecto a la repetición fallida del experimento, Newton sugiere que a Lucas le puede estar ocurriendo lo mismo que a Linus, que no replicaba los experimentos de Newton en las mismas condiciones exactas. Pero Newton es consciente de que en el caso del experimento básico del espectro, que Linus había intentado replicar, el grupo de Lieja no reconoció abiertamente que el error había sido suyo.

But when they found in your Chamber ye success in my side, they deferred to let me have any account of it till you had find out a set of new objections, ye whole stay amounting to about half a year: and then instead of acknowledging his mistake ye business was represented as if none had been committed by him" (*ibid.*, p. 259).

Tras recordarle a Lucas ese feo modo de proceder, concluye: "But if you will not yet freely let us know your mistakes but hope to mend ye matter by new disputes, you are at your liberty" (*ibid.*, p. 260). En definitiva, Newton creía que los colores primarios no estaban siendo bien separados por parte de Lucas.

Hemos visto que gran parte de la oposición a la nueva teoría de Newton procedía de una asociación del *experimentum crucis* con una concepción corpuscular de la naturaleza de la luz. Eso, junto a las escuetas descripciones del experimento proporcionadas por Newton, provocó que muchos negasen los resultados de ese experimento tras repetirlo y observar unos resultados diferentes. «Most recognized the significance of his idea that sunlight consists of diversely refrangible rays and resisted its challenge to received optical theory. Newton's principal experiments were initially questioned primarily because of his inadequate description of them» (Shapiro, 1996, p. 80). Las teorías de la luz como modificación del medio se veían amenazadas al no poder dar cuenta de por qué en el segundo prisma no se "modificaba" el rayo monocromático respecto a la primera refracción. De ese modo, aceptar que los colores eran propiedades inmutables de los rayos iba en contra de la concepción de los colores como modificaciones o desviaciones del pulso o movimiento que era la luz. En este sentido, la teoría de Newton necesitaría reunir y mostrar una cantidad de evidencia experimental mucho mayor para lograr una aceptación general.



1. LA ACEPTACIÓN GRADUAL DE LA TEORÍA DE NEWTON

A pesar de las críticas recibidas respecto a la nueva teoría de la luz y los colores y al estratégico *experimentum crucis*, que jugaba un papel fundamental a la hora de establecer que la luz era un agregado heterogéneo y consistía en rayos de diferente refrangibilidad, la nueva teoría de Newton fue ganando apoyo progresivamente. Shapiro (1996) describe cómo fue surgiendo esa aceptación gradual y achaca la resistencia inicial contra la teoría de Newton a la breve y poco clara presentación de la misma en el artículo de 1672, algo de lo que Newton fue consciente y que determinó la preparación de una obra amplia y minuciosa que no dejase lugar a dudas sobre los resultados de sus investigaciones ópticas:



www.solofici.org

By the time the *Opticks* appeared in 1704, there was substantial support in England and Scotland for Newton's theory. The key to the acceptance of the entire theory was unequal refrangibility, which had major implications for geometrical and applied optics and involved a significant extension of the law of refraction. After initial resistance, caused both by the brief, unclear presentation in the "New Theory" and also by resistance to revisions in such a major new discovery as the law of refraction, unequal refrangibility was largely uncontested among those who knew of the theory. [...] Once Newton clarified his experimental procedures, it was very easy to carry out the basic experiment of the elongated spectrum, which is the foundation for unequal refrangibility (p.131).



Desde el artículo de 1672 hasta la publicación de la *Óptica* en 1704 pasaron más de treinta años en los que la teoría de Newton recibió el apoyo de fundamentalmente dos grupos: uno de científicos matemáticos en las Islas Británicas y otro de filósofos de la naturaleza en las universidades escocesas.

La óptica geométrica era una de las principales partes de las ciencias matemáticas en la época de Newton, y la teoría de éste acerca de la diferente refrangibilidad supuso un avance para esa disciplina tanto teórico (pues ampliaba los principios en los que se basaba) como práctico (por la aplicación que tenía para la astronomía al mejorar la comprensión de la aberración cromática, producida por la refracción).



www.solofici.org

SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

Entre el grupo de los científicos matemáticos que apoyaron la teoría de Newton encontramos al astrónomo J. Flamsteed, que defendía la importancia de la diferente refrangibilidad propuesta por Newton para la comprensión de las lentes de los telescopios y de la causa de las aberraciones cromáticas. E. Halley también apoyaría la teoría newtoniana de la diferente refrangibilidad de la luz y la combinaría con una teoría de la emisión de la luz según la cual los colores se corresponderían con corpúsculos diferentes. Hemos visto que Huygens trataba a la teoría de Newton como hipótesis probable, y aunque nunca la reconoció como verdadera (pues rechazaba las implicaciones de esa teoría), sí reconoció como útil el descubrimiento de la diferente refrangibilidad de los rayos para el estudio de la aberración cromática. Así, pese a su rechazo sobre la concepción de los colores de Newton, «The *Dioptrica* was published in Huygens's *Opuscula postuma* in 1703, a year before the *Opticks*, and its endorsement of unequal refrangibility most likely paved the way for the subsequent acceptance of Newton's theory on the Continent» (*ibid.*, p. 83).

Las universidades escocesas habían abandonado los programas dominados por la filosofía de la naturaleza aristotélica y la astronomía ptolemaica, y en la década de 1670 incorporaron los principios de la nueva filosofía mecánica y experimental.



Así, a principios de la década de 1680 enseñaban ya la teoría de Newton como verdadera, aunque ésta también era combinada con la teoría cartesiana o la ondulatoria de Huygens. Esto no era extraño si tenemos en cuenta que Newton trató de señalar que las hipótesis de la luz como modificación del medio no eran incompatibles con lo que él establecía acerca de las propiedades de la luz (algo que, según él y como hemos dicho, debería aceptarse independientemente de cualquier hipótesis sobre la naturaleza de la luz): «It is a mistake to judge acceptance of Newton's theory of color by adherence to any particular conception of the nature of light and, in particular, to Newton's own emission theory» (*ibid.*, p. 86).

La aceptación de la teoría de Newton se disparará tras la publicación de la *Óptica* en 1704. En este *Tratado de las reflexiones, refracciones, inflexiones y colores de la luz*, como continúa el título de la obra, Newton ofrece una gran cantidad de experimentos descritos con gran detalle en apoyo de su teoría sobre la luz y los colores, lo cual contrasta con la escasez y brevedad del *experimentum crucis*. Así, respecto a la intención de esta obra, Shapiro (*ibid.*) argumenta:

When Newton began the *Opticks* in 1687, he had no intention of undertaking a new series of experimental investigations on his theory of color. His aim was to present his earlier investigation in a coherent and more convincing work than his earlier series of papers, which were largely responses to criticism (p. 88).

Pero si no tenía intención de realizar nuevos experimentos de cara a incorporarlos en ese nuevo tratado, ¿por qué tardó tanto en publicarlo? Shapiro (2008) analiza en qué medida la rivalidad entre Newton y Hooke influyó en ese retraso. Como ya hemos indicado, cuando Newton publica por primera vez su teoría sobre la luz y los colores en 1672, Hooke era la máxima autoridad en óptica en Inglaterra y Newton se proponía superar su programa de investigación en este campo, según el cual para una explicación de la luz había que dar cuenta de la refracción, de los colores de las láminas finas, de los colores de los cuerpos y de la difracción. Newton encontró serias dificultades para ofrecer una explicación empírica de la difracción de la luz. Había encontrado que algunos experimentos contradecían su teoría acerca de este fenómeno óptico, con lo cual esta investigación le llevó a retrasar la publicación de su *Óptica* durante unos doce años, hasta 1704. No logrando superar las dificultades que había encontrado, al final decide suprimir esta parte, que iba a constituir un cuarto libro, más allá de los tres que componen la obra. El hecho de que Newton tomase esta decisión no es muestra sino de la fidelidad a sus principios metodológicos por los cuales había que basarse en lo probado por los experimentos y evitar siempre caer en hipótesis. Pero parece que la decisión de abandonar la parte de la difracción no fue fácil para Newton, puede que debido a la rivalidad con Hooke, como opina Shapiro (*ibid.*):

Hooke had convinced him that an account of diffraction was an integral part of any complete theory of light and color, and Newton was not going to abandon this challenge even at this stage of his career. It is in this sense that I believe Hooke was responsible for the delay of the *Opticks* (p. 435).

De este modo, la muerte de Hooke podría haber sido uno de los estímulos que llevaron a Newton a decidirse a publicar la *Óptica* en 1704. Otros alicientes podrían haber sido también su elección como presidente de la Royal Society o la reciente publicación de la *Dioptrica* de Huygens en 1703. Pese a no haber incluido la parte sobre la difracción, la *Óptica* constituía un tratado muy completo sobre la luz y los colores. Ahora las proposiciones establecidas por Newton eran apoyadas por una gran cantidad de pruebas experimentales y las descripciones de las mismas eran lo suficientemente minuciosas como para no ser necesaria una replicación de todos los experimentos de cara a aceptar la teoría de Newton, cuyo núcleo era, no lo olvidemos, la diferente refrangibilidad de los rayos de luz. «There is no evidence that anyone tried systematically to repeat Newton's experiments to see whether they were true. Each chose what he judged to be critical experiments» (Shapiro, 1996, p. 95). Así, ante la gran cantidad de evidencia empírica proporcionada, la teoría de Newton fue ganando seguidores a lo largo de toda Europa. «Moreover, since the elongated spectrum was a new, essentially uncontested, phenomenon discovered by Newton, there was no long-established rival explanation to unequal refrangibility» (*ibid.*, p. 91).

SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

La diferente refrangibilidad de los rayos es, tras la publicación de la *Óptica*, progresivamente aceptada como probada. Si bien en 1672 el peso de esta proposición recaía explícitamente en un solo experimento, el *experimentum crucis*, ahora éste perdía la denominación de "crucial" y quedaba diluido entre la abrumadora colección de experimentos que Newton recogía en la *Óptica*, pasando a ser uno más. En esta obra, Newton sigue cierto formato expositivo: primero, expone los conceptos físicos que va a utilizar; luego, describe con detalle los experimentos que funcionan como pruebas empíricas de las proposiciones que pretende probar; a continuación registra lo que ha sido observado tras la realización de los experimentos; y, por último, ofrece su interpretación razonada de los resultados observados. Siguiendo este esquema y tras exponer una serie de experimentos, Newton nos ofrece el razonamiento a partir del que establece que la luz del sol es una mezcla heterogénea de rayos que constan de diferente refrangibilidad:

En todos estos casos, hemos visto que siempre hay rayos que, a incidencias iguales sobre el mismo medio, sufren refracciones desiguales, sin que se dilaten o escindan los rayos individuales o sin que se produzca una desigualdad fortuita en las refracciones, como se demostró en los experimentos quinto y sexto. Hemos visto que los rayos que difieren en refrangibilidad se pueden distinguir y separar entre sí, sea por refracción, como en el tercer experimento, sea por reflexión, como en el décimo, y que los diversos tipos, una vez separados, sufren refracciones desiguales a incidencias iguales. También hemos visto que, tras la separación, se refractan más aquellos rayos que más se refractaban antes de la misma, como se muestra en el experimento sexto y siguientes, y que si se hace que la luz del Sol atraviese sucesivamente tres o más prismas cruzados, los rayos que más se refractaban en el primer prisma se refractan más que los otros, y en la misma razón y proporción, en los prismas siguientes, como se desprende del quinto experimento. Por todas estas consideraciones, es manifiesto que la luz del Sol es una mezcla heterogénea de rayos, unos de los cuales son constantemente más refrangibles que los otros, como habíamos propuesto (Newton, 1977, p. 62).



Aquí, Newton establece que la luz consta de rayos que difieren en refrangibilidad, lo cual es distinto de establecer que la luz consta de rayos que difieren en color, pero una vez aceptada la primera proposición, se allanaba el camino para aceptar la segunda:

«By clearly isolating unequal refrangibility in the first part and presenting it in great detail, Newton structured the text so that the more difficult and paradoxical part on color followed almost as a natural consequence» (Shapiro, 1996, p. 91). Diez años más tarde de la publicación de la *Óptica*, J. T. Desaguliers, asistente y divulgador de Newton, lleva a cabo una demostración de la inmutabilidad del color que será clave para una aceptación íntegra de la teoría de Newton. La importancia de esta demostración radicaba en que las teorías modificacionistas se veían rebatidas por esa inmutabilidad del color: según estas teorías, el color dependía siempre de una modificación de la luz, fuese por refracción o por reflexión.

Además, la demostración de Desaguliers se hacía necesaria porque la inmutabilidad del color había sido desafiada por E. Mariotte en Francia, donde las teorías cartesianas y neo-cartesianas estaban muy arraigadas. Un año después de la publicación del artículo de Newton de 1672, Mariotte había empezado a desarrollar una serie de pruebas de cara a desafiar el *experimentum crucis* y demostrar que la teoría de la inmutabilidad del color era falsa. Mariotte defendía que el color de un rayo de luz podía ser cambiado en la refracción. Su supuesta refutación de la teoría de Newton tuvo mucha trascendencia en el continente europeo. Pero, como describe Schaffer (1989): «In his version of Newton's experiment with two prisms, he did not place the first prism at minimum deviation, nor did he place a screen immediately after this prism to collimate the rays produced» (p. 93). Para que la *Óptica* fuese aceptada en Francia era necesario dejar claro que la refutación de Mariotte no era válida. Newton redacta la introducción al artículo de Desaguliers y en él asegura que si Mariotte no halló que los colores fuesen inmutables tras repetir el *experimentum crucis*, fue porque no estaba al tanto de la forma correcta en la que había que separarlos y de los detalles necesarios para usar los instrumentos correctos, ya que esto no fue publicado hasta la *Óptica*.

The manner of separating the primitive Colours of Light to such a Degree, that if any one of the separated Lights be taken apart, its Colour shall be found unchangeable, was not published before Sir Is. Newtons Opticks came abroad. For want of knowing how this was to be done, some Gentlemen of the English College at Liege, and Monsieur Mariotte in France, and some others took those for primitive Colours, which are made by immitting a Beam of the Sun's Light into a dark Room through a small round Hole, and refracting the Beam by a triangular Prism of Glass placed at the Hole. And by trying the Experiment in this manner, they found that the Colours thus made were capable of change, and thereupon reported that the Experiment did not succeed [parte de la introducción escrita por Newton] (Newton y Desaguliers, 1714, p. 433).

SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

Así, en cierto modo, Newton reconoce que el *experimentum crucis* tal como fue presentado en su artículo de 1672 no produce una separación perfecta de los colores.

The *experimentum crucis* and, more generally, ordinary refraction in a single prism are not capable of producing rays of a single color. To understand why this was so, Newton explained [...] that the spectrum was composed of overlapping circular images of the sun, each of a different degree of refrangibility and color. The way to produce a spectrum of simple colors is obvious from this construction, namely, reduce the size of the circles, or the sun's apparent diameter, so that they overlap less [...]. Newton proposed two ways to accomplish this: either insert a lens after the prism or reduce the size of the holes (Shapiro, 1996, p. 109).

Teniendo esto en cuenta, mientras en la *Óptica* el antiguo *experimentum crucis* (Libro I, Parte I, Proposición 2, Experimento 6) contribuye, junto a otros experimentos, a probar que la luz consiste en rayos de diferente refrangibilidad, Newton distingue esta prueba de la de "separar los rayos heterogéneos de la luz compuesta". Para este último propósito, Newton describe un experimento diferente al *experimentum crucis*, al que se refiere como: "Utilización de una lente para proyectar nítidamente las imágenes, así como de agujeros de diversa forma. Detalles que deben ser cuidados para el mayor éxito de la separación de los rayos"⁸.



No obstante, no es que Newton hubiese ideado este experimento para combatir las críticas a la inmutabilidad del color defendida en el artículo de 1672: es que por aquel entonces ya había realizado este experimento, como se recoge en sus *Optical Lectures* (no publicadas), pero no lo había incluido en la escueta presentación de su "New Theory". En este sentido, la nueva demostración de este experimento, entre otros, por parte de Desaguliers, junto a la traducción francesa de la *Óptica* ayudarán a extender la aceptación de la teoría de Newton en Francia.

En el artículo de Desaguliers (cuya introducción había sido redactada por Newton, como hemos visto), se presentaban nueve experimentos. El objetivo era pasar de demostrar la diferente refrangibilidad de los rayos a demostrar la inmutabilidad del color, proporcionando una descripción de los instrumentos utilizados y de cómo había que disponer el experimento. Así, el "Experimento VIII" constituiría una réplica del experimento II de la *Óptica*, por el que Newton demostraba cómo se podía lograr un espectro de rayos homogéneos.

8. Este experimento es el "Experimento II" de la *Óptica*, que prueba la Proposición IV de la Parte I del Libro I (Newton, 1977, p. 66).

En este experimento, la luz del sol atraviesa primero un pequeño agujero para ser reflejada por un espejo. Desde este espejo, atraviesa otro agujero y luego pasa por una lente antes de ser refractada por el prisma. Los rayos que salen refractados de ese prisma formarían un espectro en la pared en el que los rayos aparecen con colores mucho mejor separados:

The Colours in this Case were very vivid and well separated, only the Violet had some pale Light darting from its End, upon account of some Veins in the Prism A, and the Light not coming directly from the Sun, but reflected; which ought not to have been, if the Sun had been low enough to have thrown the Rays a good way into the Room without the Help of a Looking Glass (Newton y Desaguliers, 1714, p. 445).

El espejo ayudaba a controlar la dirección de la luz que provenía del sol. El segundo agujero sería el único cambio introducido por Desaguliers respecto al experimento II de la *Óptica* de Newton, pero en la introducción al artículo éste ya advertía de algunas precisiones a tener en cuenta para que el resultado del experimento fuese exitoso. Desaguliers termina su demostración de la inmutabilidad del color con un último experimento, «Tho shew that the Colours in this Spectrum were simple and homogeneal Lights, I made the following Experiments» (*ibid.*, p. 446). Este experimento también estaba en la *Óptica*⁹. Desaguliers procede a hacer un agujero en el panel en el que se reflejaba el espectro en el experimento anterior, de manera que cuando cualquier color pasaba a través de ese agujero, la imagen obtenida en un panel ulterior era ese color homogéneo: «I made the Experiment upon all Colours, which by this means appear'd to be simple and homogeneal» (*ibid.*).

Aunque las demostraciones de Desaguliers serían de gran importancia en Francia para la aceptación de la teoría de Newton, en los demás países, esa aceptación también dio un salto cualitativo tras la publicación de la *Óptica* en 1704. Por aquel entonces, los *Principia* ya habían determinado la orientación de la ciencia británica. Así, las concepciones físicas de Newton, su teoría corpuscular de la materia y su concepción de las fuerzas de atracción contaban con múltiples seguidores en las islas inglesas. También lo hizo su teoría corpuscular de la luz, pese a que ésta no tenía cabida en el desarrollo de la *Óptica* y era relegada a las "Cuestiones" de la misma, como ya hemos visto.

9. Este experimento es el "Experimento 12" de la *Óptica*: "El espectro producido por un prisma sobre el que incide un rayo de luz procedente de una región del espectro producido por otro, es de forma redonda". Este experimento contribuye a probar la Proposición V de la Parte I del Libro I, que reza así: "La luz homogénea se refracta regularmente, sin que los rayos se dilaten, separen o dispersen, debiéndose la visión confusa de los objetos observados a través de cuerpos refractantes, cuando se emplea luz heterogénea, al hecho de que los distintos tipos de rayos poseen diferente refrangibilidad". (Newton, 1977, pp. 70-71).

SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

La metodología de Newton también era respaldada por muchos, aunque la idea de que la filosofía natural debía estar basada en experimentos no era nueva. Estos factores también determinaron la expansión de la óptica de Newton por los países del continente. Al mismo tiempo, la teoría de la luz como modificación del medio iba perdiendo aliados. Shapiro (1996) explica así el declive de la misma:

In the 1670s all of Newton's critics supported some sort of modification theory, whereas in 1704 he encountered relatively little opposition of this sort. Modification theories of color, the only alternative to Newton's, were by the turn of the century in serious decline. This decline was caused by the decline of the most venerable of all modification theories, the Aristotelian, in which colors were produced from a mixture of blackness and whiteness, or light and shadow. This change, too, began with artists, who held that a mixture of white and black can produce nothing but grays, not chromatic colors (p. 98).

12. CONCLUSIÓN. CONSIDERACIONES DESDE LA FILOSOFÍA DE LA CIENCIA POSTERIOR SOBRE LOS EXPERIMENTOS CRUCIALES.



A lo largo del trabajo, hemos intentado proporcionar una comprensión del que, en un momento determinado, fue el *experimentum crucis* de Newton, partiendo de una presentación general del contexto teórico a partir del cual ese experimento adopta el papel de "crucial" para decidir entre dos teorías alternativas que parecían incompatibles entre sí. Así, tras la descripción de la presentación inicial del experimento y de su pretensión demostrativa dentro del artículo de 1672, hemos pasado a analizar las controversias suscitadas en los rivales de Newton, que se negaban a ver en ese experimento una prueba determinante para rebatir las teorías de la luz como modificación, por lo que no dudaron en considerar la nueva teoría de Newton como nada más que una hipótesis, a lo sumo probable, pero no verdadera. También hemos reportado brevemente cómo en la posterior aceptación de la teoría de Newton jugó un papel decisivo la argumentación presentada en la *Óptica*, obra en la que el *experimentum crucis* pasó a ser otro de los tantos experimentos que en su conjunto soportaban el peso de probar una teoría que agrupaba un gran número de proposiciones diferentes.

Respecto al papel crucial que en su momento fue atribuido a tal experimento, es interesante analizar la idea de P. Duhem según la cual los experimentos cruciales no son posibles. Para Duhem (1976) existe una interdependencia fundamental entre teoría y experimentación: «it is impossible to leave outside the laboratory door the theory that we wish to test, for without theory it is impossible to regulate a single instrument or to interpret a single reading» (p. 3).

Desde esta concepción, considera que un experimento físico no puede refutar una hipótesis aislada, como mucho, puede cuestionar un sistema o grupo de hipótesis que conforman una teoría, mostrando que entre ellas hay al menos una que no se sostiene con lo que muestra el experimento. Por lo tanto, un experimento crucial que nos permitiese decidirnos por una teoría y refutar las rivales no sería posible en física, puesto que habría que explicitar todas las hipótesis intrincadas en un sistema concreto, lo cual es, en opinión de Duhem, quimérico.

Unlike the reduction to absurdity employed by geometers, experimental contradiction does not have the power to transform a physical hypothesis into an indisputable truth; in order to confer this power on it, it would be necessary to enumerate completely the various hypotheses which may cover a determinate group of phenomena; but the physicist is never sure he has exhausted all the imaginable assumptions. The truth of a physical theory is not decided by heads or tails (*ibid.*, p. 6).

En el caso del *experimentum crucis*, tanto la teoría de Newton como las rivales estaban permeadas por una cantidad de premisas o asunciones teóricas que influían en el modo de interpretar dicho experimento. Por ejemplo, en el controvertido caso de la proposición que establecían una correspondencia unívoca entre colores y rayos de refrangibilidad, la concepción de los rayos de luz, como partes separables e independientes que componían la luz, jugaba un papel importante. El concepto de rayo, sin embargo, era un concepto teórico que Newton, a pesar de que lo admitía como "premisa", negaba que se tratase de una entidad hipotética. Las premisas admitidas por Newton, asumidas sin ofrecer una justificación de las mismas, condicionan no solo todas sus proposiciones teóricas, sino también la interpretación misma de los experimentos en los que se basan esas proposiciones. Al igual que había hecho en su carta a Pardies, en la que explicitaba las definiciones de la luz y los rayos que él tomaba como premisas, al principio de la *Óptica*, Newton también proporcionaba una serie de definiciones, de las que partía para, a continuación, poder explicar las propiedades de la luz no por hipótesis sino mediante la razón y los experimentos.

Duhem dice que esto es lo normal: siempre hay conceptos implícitos que se asumen desde una determinada concepción física, y si bien es posible también refutar alguna de esas premisas o asunciones auxiliares, para hacerlo siempre hemos de basarnos en otras nuevas asunciones. Esto es un callejón sin salida:

To seek to separate each of the hypotheses of theoretical physics from the other assumptions on which this science rests in order to subject it in isolation to observational test is to pursue a chimera; for the realization and interpretation of no matter what experiment in physics imply adherence to a whole set of theoretical propositions (*ibid.*, p. 21).

SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

Sabemos que en física siempre utilizamos ciertas definiciones o premisas que asumimos como ciertas y no esperamos que vayan a ser contradichas por experimentos. Pero, como asegura Duhem, tenemos ejemplos de casos en los que este tipo de definiciones han dejado de ser asumidas como ciertas para poder avanzar. Por citar uno de esos ejemplos, esto ha ocurrido con la propagación rectilínea de la luz, una asunción que se daba por cierta y fue rechazada para poder dar cuenta del problema de la difracción observado por Grimaldi. De este modo, los fundamentos de la ciencia a veces sufren cambios para que el progreso sea posible. Sin embargo, en ciencia siempre seguimos adoptando otro tipo de premisas o definiciones fundamentales que no se derivan de los experimentos: estas asunciones están más allá de una refutación directa por experimentos, pero esto no significa que estén a salvo de ser algún día refutadas.

On that day some one of our hypotheses, which taken in isolation defied direct experimental refutation, will crumble with the system it supported under the weight of the contradictions inflicted by reality on the consequences of this system taken as a whole (*ibíd.*, p. 37).

www.solofici.org



Mientras la postura de Duhem en contra del supuesto rol de “crucialidad” que se ha atribuido a algunos experimentos es firme, otros autores, como Alejandro Cassini, reivindican la existencia de dichos experimentos cruciales, si bien entendidos de una forma más laxa. Así, Cassini (2015) cree que los experimentos cruciales son posibles en ciencia, aunque, como dice Duhem, no refuten una teoría o hipótesis determinada sino un conjunto heterogéneo de hipótesis sin decirnos dónde está el error. Para Cassini, los experimentos cruciales no necesariamente tienen por qué ser decisivos, esto es, proporcionar razones suficientes para aceptar una de las teorías y rechazar las otras (esto es lo que se entendería desde una concepción “fuerte” de un experimento crucial), por lo que cree que siempre pueden describirse los resultados de un experimento considerado crucial independientemente de las teorías rivales en disputa. Cassini, por lo tanto, rechaza una concepción fuerte de los experimentos cruciales según la cual estos refutarían o verificarían hipótesis rivales. Desde el punto de vista de este autor, esto solo podría darse entre dos hipótesis rivales que fuesen además contradictorias entre sí, es decir, que la refutación de una implicaría la verificación de la otra y no sería posible imaginar una tercera hipótesis alternativa. «En cambio, si las dos hipótesis rivales no son contradictorias, de la refutación de una de ellas nada se sigue acerca de la verificación de la otra, ya que podrían ser ambas falsas» (Cassini, 2015, p. 108). Cassini, por lo tanto, defiende una interpretación más débil de los experimentos cruciales, admitiendo la interdependencia existente entre teoría y experimentación:

Cualquier resultado experimental se interpreta en el contexto de un amplio sistema de hipótesis y teorías admitidas, esto es, cualquier experimento está cargado de teoría. [...]

Así, aun admitiendo un holismo epistemológico moderado (que no implique que en cada experimento se contrasta la totalidad de nuestro conocimiento), [...], los experimentos cruciales resultan perfectamente posibles (*ibíd.*, p. 133).



Esta interpretación más débil de la crucialidad de los experimentos no es por ello menos válida y pasa irremediablemente por la comprensión del experimento crucial en cuestión respecto al contexto teórico a partir del que surgió y a la repercusión inmediata que pudo tener en su ámbito científico.

Los experimentos cruciales no son “cruciales en retrospectiva”, sino casi siempre en el momento mismo en que se los realiza. Generalmente, los experimentos cruciales influyen de manera significativa en el curso del desarrollo futuro de las teorías. [...] Hace ya mucho tiempo que los filósofos de la ciencia han aceptado que ningún experimento, desde un punto de vista estrictamente lógico, es capaz de verificar o refutar una teoría de manera concluyente. ¿Por qué la situación debería ser diferente con los experimentos cruciales? La evidencia proporcionada por un experimento crucial, como la de cualquier otro experimento, sólo puede pretender confirmar o disconfirmar una teoría o sistema de hipótesis dentro de un contexto teórico específicamente fijado, pero siempre susceptible de revisión. Eso es lo más que puede esperarse de un resultado experimental en general (*ibíd.*, p. 134).

En este trabajo hemos intentado ofrecer una comprensión del *experimentum crucis* de Newton desde esta perspectiva, presentando el contexto teórico en el que ese experimento llegó a tener sentido como “experimento crucial” y analizando la repercusión que tuvo en el momento histórico de su presentación. Aunque, como dice Duhem, los experimentos cruciales definitivos no sean posibles en ciencia (pues no existen ni las refutaciones ni las confirmaciones definitivas), eso no quiere decir que no haya experimentos cruciales que en un determinado momento confirman una teoría y refutan otra. En este sentido, el *experimentum crucis* de Newton tuvo una gran importancia para la demostración de su teoría de la luz y los colores y logró poner entre la espada y la pared a las teorías modificacionistas tal y como estaban desarrolladas en aquel momento.



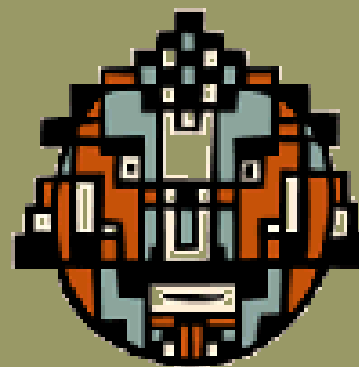
SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

13. BIBLIOGRAFÍA

- Birch, Thomas (1757). *The History of the Royal Society of London, for improving of Natural Knowledge, from its first rise*, Vol. III. Londres: A. Millar.
- Cassini, Alejandro (2015). "Una reivindicación de los experimentos cruciales". *Revista de Filosofía*, 40, 1: 105-137.
- Descartes, René (1986). *Discurso del Método, Dióptrica, Meteoros y Geometría*. Madrid: Alfaguara.
- Descartes, René (1989). *El mundo. Tratado de la luz*. Barcelona: Anthropos.
- Duhem, Pierre (1976). "Physical Theory and Experiment". En Sandra G. Harding (Ed.), *Can Theories be Refuted? Essays on the Duhem-Quine Thesis*. Holanda: D. Reidel Publishing Company: 1-40.
- Gascoines, John (1959). "148 Gascoines to Oldengurg. 15 December 1675". En Turnbull, H. W. (ed.), *The Correspondence of Isaac Newton*. Vol I (1661-1675). Cambridge (Published for the Royal Society) at the University Press: 393-396.
- Hooke, Robert (1989). *Micrografía o algunas descripciones fisiológicas de los cuerpos diminutos realizadas mediante cristales de aumento con observaciones y disquisiciones sobre ellas*. Madrid: Alfaguara.
- Hooke, Robert (1672). "Robert Hooke's Critique of Newton's Theory of Light and Colors". En Birch, Thomas (1757), *The History of the Royal Society*. London: 3: 10-15.
- Huygens, Christiaan (1673). "An Extract of a Letter lately written by an ingenious person from Paris [Christiaan Huygens]". *Philosophical Transactions of the Royal Society*, 8, 96: 6086-6087.
- Lucas, Anthony (1959). "161 Lucas to Oldenburg. 16 May 1676". En Turnbull, H. W. (ed.), *The Correspondence of Isaac Newton*. Vol II (1676-1687). Cambridge (Published for the Royal Society) at the University Press: 8-14.
- Lucas, Anthony (1959). "185 Lucas to Oldenburg. 13 October 1676". En Turnbull, H. W. (ed.), *The Correspondence of Isaac Newton*. Vol II (1676-1687). Cambridge (Published for the Royal Society) at the University Press: 104-108.
- McGuire, J. E. (1977). "Neoplatonism and Active Principles: Newton and the Corpus Hermeticum". En R. S. Westman y J. McGuire, *Hermeticism and the Scientific Revolution*. Los Angeles: University of California: 94-142
- Newton, Isaac (1959). "143 Newton to Oldenburg. 13 November 1675". En Turnbull, H. W. (ed.), *The Correspondence of Isaac Newton*. Vol. I (1661-1675). Cambridge (Published for the Royal Society) at the University Press: 356-359.
- Newton, Isaac (1959). "220 Newton to Lucas. 5 March 1677/8". En Turnbull, H. W. (ed.), *The Correspondence of Isaac Newton*. Vol. II (1676-1687). Cambridge (Published for the Royal Society) at the University Press: 254-261.
- Newton, Isaac (1671/2). "A Letter of Mr. Isaac Newton ... containing his New Theory about Light and Colors". *Philosophical Transactions of the Royal Society*, 6, 80: 3075-3087.
- Newton, Isaac (1672). "A Serie's of Quere's propounded by Mr. Isaac Newton ... positively concluding his new Theory of Light and Colours". *Philosophical Transactions of the Royal Society*, 7, 85: 5004-5007.
- Newton, Isaac (1671/2). "Draft of 'A Theory Concerning Light and Colors'". MS Add. 3970.3. Cambridge: Cambridge University Library: 460-466.
- Newton, Isaac (1672). "Mr Isaac Newtons Answer to some Considerations [of Robert Hooke] upon his doctrine of Light and Colors". *Philosophical Transactions of the Royal Society*, 7, 88: 5084-5103.
- Newton, Isaac (1673). "Mr. Newtons Answer to the foregoing Letter [of Christiaan Huygens]". *Philosophical Transactions of the Royal Society*, 8, 96: 6087-6092.
- Newton, Isaac (1958). "Mr. Newtons Answer to the foregoing Letter" [Respuesta a la segunda carta de Pardies]. (Versión original en latín: 1672, *Phil. Trans.*, 7, 85: 5014-5018). En Cohen, I. Bernard (ed.), *Isaac Newton's Papers and Letters on Natural Philosophy and Related Documents*. Cambridge University Press: 106-109.
- Newton, Isaac (1958). "Mr. Newtons Letter ... being an Answer to the foregoing Letter of P. Pardies" [Respuesta a la primera carta de Pardies]. (Versión original en latín: 1672, *Phil. Trans.*, 7, 84: 4091-4093). En Cohen, I. Bernard (ed.), *Isaac Newton's Papers and Letters on Natural Philosophy and Related Documents*. Cambridge University Press: 83-85.
- Newton, Isaac (1977). *Óptica*. Madrid: Alfaguara.
- Newton, Isaac (1987). *Principios matemáticos de la filosofía natural*, 1 y 2. Madrid: Alianza Editorial.
- Newton, Isaac y Desaguliers, J. T. (1714). "An Account of Some Experiments of Light and Colours, Formerly Made by Sir Isaac Newton, and Mention'd in His Opticks, Lately Repeated before the Royal Society by J. T. Desaguliers, F. R. S.". *Philosophical Transactions of the Royal Society (1714-1716)*, 29: 433-447.

SEGUNDO PREMIO: El *experimentum crucis* de Newton. Contexto teórico y controversias suscitadas

- Pardies, Ignatius Gaston (1958). "A Latin Letter ... by Ignatius Gaston Pardies ... containing some Animadversions upon Mr. Isaac Newton ... his Theory of Light ...". (Versión original en latín: 1672, *Phil. Trans.*, 7, 84: 4087-4090). En Cohen, I. Bernard (ed.), *Isaac Newton's Papers and Letters on Natural Philosophy and Related Documents*. Cambridge University Press: 79-82.
- Pardies, Ignatius Gaston (1958). "A Second Letter of P. Pardies ... to Mr. Newtons Answer, made to his first Letter ...". (Versión original en latín: 1672, *Phil. Trans.*, 7, 85: 5012-5013). En Cohen, I. Bernard (ed.), *Isaac Newton's Papers and Letters on Natural Philosophy and Related Documents*. Cambridge University Press: 104-105.
- Raftopoulos, Athanassios (1999). "Newton's Experimental Proofs as Eliminative Reasoning". *Erkenntnis*, 50, 1: 95-125.
- Sabra, Abdelhamid I. (1981). *Theories of Light from Descartes to Newton*. Cambridge: Cambridge University Press.
- Schaffer, Simon (1989). "Glass works: Newton's prisms and the uses of experiment". En David Gooding, Trevor Pinch y Simon Schaffer (Eds.), *The uses of experiment: studies in the natural sciences*. Cambridge: Cambridge University Press: 67-104.
- Shapiro, Alan E. (2004). "Newton's optics and atomism". En I. Bernard Cohen y George E. Smith (Eds.), *The Cambridge Companion to Newton*. Cambridge: Cambridge University Press: 227-255.
- Shapiro, Alan E. (1979). "Newton's "Achromatic" Dispersion Law: Theoretical Background and Experimental Evidence". *Archive for History of Exact Sciences*, 21, 2: 91-128.
- Shapiro, Alan E. (1996). "The Gradual Acceptance of Newton's Theory of Light and Color, 1672-1727". *Perspectives on Science*, 4, 1: 59-139.
- Shapiro, Alan E. (2008). "Twenty-Nine Years in the Making: Newton's *Opticks*". *Perspectives on Science*, 16, 4: 417-438.
- Westfall, Richard S. (1980). *Never at Rest. A Biography of Isaac Newton*. Cambridge: Cambridge University Press.
- Worrall, John (2000). "The Scope, Limits, and Distinctiveness of the Method of 'Deduction from the Phenomena': Some Lessons from Newton's 'Demonstrations' in Optics". *The British Journal for the Philosophy of Science*, 51, 1: 45-80.



SOCIEDAD DE
LÓGICA,
METODOLOGÍA Y
FILOSOFÍA DE LA
CIENCIA EN
ESPAÑA

www.solofici.org



TERCER PREMIO: Inferencialismo en el discurso ético

Llanos Navarro Laespada

Directora: María José Frápolli Sanz
Máster en Lógica y Filosofía de la Ciencia-UGR

ÍNDICE DE CONTENIDOS

Agradecimientos.....	i
Resumen/Abstract.....	1
Palabras clave/Keywords.....	2
1. Introducción.....	2
2. Moore y el argumento de la pregunta abierta.....	7
3. Frege y el argumento de la pregunta abierta. Una explicación no representacionalista de las atribuciones de verdad.....	9
4. Una explicación no representacionalista de las atribuciones de bondad.....	5
5. Funciones de predicados y el problema Frege-Geach.....	20
6. El anti-representacionalismo como tesis global.....	22
7. Conclusiones.....	26
Referencias bibliográficas	



AGRADECIMIENTOS

Agradezco a María José Frápolli que me haya acompañado en la elaboración de este trabajo. Su dirección ha hecho de esta investigación un proceso estimulante y enriquecedor. Gracias.

RESUMEN



El expresivismo ético es una teoría del significado de las oraciones y los términos éticos que trata de responder a los problemas que la concepción representacionalista del significado encuentra en este tipo de expresiones. El expresivismo mantiene que las afirmaciones éticas no representan estados de cosas, sino que expresan estados conativos o actitudes del hablante, y que en virtud de ello son significativas. El hecho de que las oraciones éticas actúan como argumentos de operadores veritativo-funcionales ha obligado al expresivismo ético a admitir, en contra de sus primeros partidarios, que oraciones éticas como “Compartir es bueno” tienen contenido proposicional. No obstante, la teoría ha tenido dificultades para hacer consistente esta idea y defender, al mismo tiempo, la especificidad del significado de este tipo de oraciones.

El presente trabajo propone entender los términos éticos como funciones de orden superior, en el contexto de una teoría inferencialista del significado. Ello permitirá, por un lado, dar cuenta del contenido proposicional de las oraciones éticas y, por otro, recoger la intuición expresivista de que el tipo de acción que un hablante realiza cuando hace una afirmación ética no consiste en una representación del mundo.

Abstract

Ethical expressivism is a theory of the meaning of ethical sentences and terms which aims to respond to the problems that the descriptivist conception of meaning finds in this type of expressions. Expressivism holds that ethical claims do not represent states of affairs, but express conative states or speaker's attitudes, and that they are meaningful in virtue of that. The fact that ethical sentences act as arguments of truth-functional operators has forced ethical expressivism to admit, against its first supporters, that ethical sentences such as “Sharing is good” have propositional content. Nevertheless, the theory has had difficulty in making such idea consistent and defending, at the same time, the specificity of the meaning of this type of sentences. The present paper proposes to understand ethical terms as higher-order functions, in the context of an inferentialist theory of meaning. That will allow, on one hand, for an account of the propositional content of ethical sentences and, on the other hand, for preserving the expressivist intuition in which the type of action carried out by a speaker when making an ethical claim does not consist of a representation of the world.



www.solofici.org

Palabras clave

Funciones de orden superior, expresivismo global, modelo orgánico, naturalismo ontológico.

Keywords

Higher-order functions, global expressivism, organic model, ontological naturalism.

TERCER PREMIO: Inferencialismo en el discurso ético

I. Introducción



El expresivismo ético surge en el siglo XX como una teoría del significado de las oraciones y los términos éticos con el objetivo de responder a los problemas que la concepción descriptivista o representacionista del significado encuentra en este tipo de expresiones. El representacionismo sostiene que el significado de una oración consiste en la representación de un estado de cosas. De acuerdo con ello, la oración “La nieve es blanca” es significativa porque representa un estado de cosas en el que la referencia del término “nieve” tiene la propiedad a la que refiere el predicado “blanca” y, del mismo modo, una oración como “Es bueno compartir” es significativa porque representa un estado de cosas en el que la referencia de “compartir” tiene la propiedad designada por el predicado “bueno”. Es decir, el representacionismo mantiene que el significado de las oraciones éticas depende de ciertas partes de la realidad que estas describen, en particular, de las propiedades éticas a las que supuestamente refieren los predicados éticos (véanse Chrisman, 2011: 29; Frápolli y Villanueva, 2013: 584). Pero la idea de que el significado del discurso ético depende de propiedades éticas tiene que enfrentarse con el siguiente problema: explicar la motivación de los hablantes a actuar de acuerdo con sus juicios éticos. A diferencia de creencias como la creencia de que la nieve es blanca, las supuestas creencias éticas parecen ser capaces de motivar la conducta por sí mismas, sin necesidad de asociarse con un estado conativo (véase Chrisman, 2011: 32; Stevenson, 1937: 16). En consecuencia, las propiedades de las que, según el representacionismo, dependería el significado de las afirmaciones y las creencias éticas han de ser *sui generis*: a diferencia de las propiedades naturales, las propiedades éticas tendrían la capacidad de motivar por sí solas la conducta de quien las percibe. A fin de evitar compromisos ontológicos anti-naturalistas, el representacionista podría mantener que estas propiedades, aunque *sui generis*, son reductibles a propiedades naturales. Sin embargo, hasta el momento no ha sido posible explicar cómo se produciría esta reducción, ni a través de una explicación metafísica de la naturaleza de las propiedades éticas, ni a través de un análisis de los conceptos éticos (Chrisman, 2011: 35). Moore (1903/2002) ha argumentado que los predicados éticos son irreducibles a predicados descriptivos, ya que dada cualquier definición de “bueno” en términos de propiedades naturales es posible preguntar con sentido si un objeto que tiene esas propiedades es bueno. A menos que logre explicar de qué modo se efectuaría la reducción, el representacionista se ve obligado a escoger entre dos opciones: o bien renuncia a una ontología naturalista, o bien asume una teoría del error respecto del discurso ético.



El expresivismo ético evita el dilema mediante una explicación anti-representacionista del significado de las oraciones éticas. Los expresivistas aceptan, con Moore, que los conceptos éticos no pueden ser reducidos a conceptos descriptivos (sirvan como ejemplo Ayer, 1936/1971: 106-107; Stevenson, 1937: 18), pero se resisten a postular la existencia de propiedades no naturales. Su estrategia consiste en cambiar la pregunta: para comprender el significado de las oraciones éticas no debemos preguntar a qué propiedades refieren los términos éticos, sino en qué consisten las evaluaciones éticas, esto es, qué tipo de acción realizan los hablantes cuando hacen afirmaciones éticas (véanse Chrisman, 2011: 30; Gibbard, 2003/2009: 6). En respuesta a esta pregunta los primeros expresivistas, o emotivistas, admitieron la que Price –siguiendo a Rorty– ha llamado “tesis de la bifurcación” (Price, 2013: 30). La tesis de la bifurcación establece un contraste entre dos clases de afirmaciones que serían significativas por razones diferentes. Por un lado, los emotivistas conservaron la explicación representacionista para las afirmaciones que no involucran términos éticos y, por otro, sostuvieron que el significado de las afirmaciones éticas consiste en la expresión de estados mentales conativos. Así, Stevenson afirma que hay dos propósitos que llevan a los hablantes a emplear el lenguaje: el propósito de registrar, aclarar y comunicar creencias y el propósito de manifestar nuestros sentimientos, crear estados de ánimo o incitar a las personas a acciones o actitudes (Stevenson, 1937: 21). Según dice, los términos éticos responden al segundo de esos propósitos; «su uso fundamental no es indicar hechos, sino crear influencias» (Stevenson, 1937: 18). Por su parte, Ayer afirma que el significado de las afirmaciones éticas debe ser definido en términos de las emociones que estas expresan y de las respuestas que pretenden provocar (Ayer, 1936/1971: 111). Según él, los términos éticos tienen la función no descriptiva de expresar o evidenciar emociones del hablante y, por ello, no contribuyen al contenido factual de las oraciones en las que aparecen (Ayer, 1936/1971: 110). Una consecuencia que Ayer extrae de la tesis de que los términos éticos no describen es que hay afirmaciones éticas que, aunque son significativas, no son susceptibles de ser verdaderas o falsas. Un ejemplo lo constituiría la afirmación “Robar está mal”. De acuerdo con Ayer, lo que un hablante hace por medio de esa afirmación es mostrar su desaprobación ante la acción de robar, de modo que la afirmación sería equivalente a la preferencia de “robar” con un cierto tono de desagrado o rechazo (Ayer, 1936/1971: 110).

En “Ascriptivism” (1960), Geach propone un argumento contra las posiciones no descriptivistas que ha pasado a ser conocido como el “argumento Frege-Geach” y que se ha convertido en la crítica más exitosa en contra del emotivismo. Según él, las teorías no descriptivistas del significado ignoran de manera indebida la diferencia que

TERCER PREMIO: Inferencialismo en el discurso ético

Frege establece entre considerar un contenido y afirmarlo. Más concretamente, estas teorías no tienen en cuenta que las oraciones éticas pueden caer bajo el alcance de conectivas lógicas y que, en esos usos, esas oraciones no son afirmadas. Según Geach, si el emotivista estuviera en lo cierto, entonces un argumento como “Si X es bueno, entonces p. X es bueno. Luego, p” contendría una falacia de equivocidad. La segunda premisa debería su significado a la acción que el hablante realiza cuando afirma esa oración, mientras que el significado de la misma oración en la primera premisa debe ser explicado de otro modo. Es decir, la equivocidad radicaría en que el significado del antecedente de la primera premisa y el significado de la segunda premisa serían distintos, a pesar de que ambas oraciones son idénticas desde el punto de vista lingüístico. Pero, a juicio de Geach, no hay tal equivocidad: el argumento es claramente válido. En consecuencia, el emotivista se equivoca (Geach, 1960: 223).

El argumento de Geach pone de manifiesto que las oraciones éticas pueden caer bajo el alcance de operadores veritativo-funcionales. Este hecho sugiere, en contra del emotivismo, que estas oraciones son susceptibles de ser verdaderas o falsas (Frápolly y Villanueva, 2012: 477; 2013: 592). El emotivismo ofrece una explicación de la especificidad de los predicados éticos, pero no puede dar cuenta del hecho de que las oraciones que los involucran se comportan como si tuvieran valor de verdad. Esto ha motivado el surgimiento de otras propuestas expresivistas. En general, estas nuevas propuestas conservan la tesis emotivista de que las afirmaciones éticas tienen la función no descriptiva de expresar una actitud del hablante, pero no renuncian a reconocer contenido proposicional en estas afirmaciones. Son ejemplos de ello el cuasi-realismo de Blackburn (1993), el expresivismo ecuménico de Ridge (2006) y el neo-expresivismo de Bar-On y Chrisman (2009).

Blackburn persiste en mantener que las afirmaciones éticas no son susceptibles de ser verdaderas o falsas. Su solución al problema Frege-Geach consiste en afirmar que las oraciones que no tienen valor de verdad también pueden formar parte de argumentos. Consideremos el condicional “Si mentir es malo, entonces incitar a tu hermano a mentir es malo”. Según Blackburn, afirmar ese condicional equivale a expresar la aprobación de las sensibilidades éticas que afirman “Incitar a tu hermano a mentir es malo” cuando afirman “Mentir es malo”, es decir, las sensibilidades que desapruaban que alguien incite a su hermano a mentir cuando desapruaban mentir (Blackburn, 1993: 188; véase Miller, 2003: 59). A pesar de que esas afirmaciones no tienen valor de verdad, dice Blackburn, sería incoherente que un hablante afirmara el condicional y su antecedente pero negara el consecuente. La incoherencia se debe a que esas preferencias expresan objetivos que no pueden ser satisfechos conjuntamente es imposible que las actitudes que el hablante expresa al afirmar las premisas (Blackburn, 1993: 189); es imposible que las actitudes que el hablante expresa al afirmar las premisas sean satisfechas a la vez que es satisfecha la

actitud que el hablante expresa cuando niega la conclusión. En definitiva, Blackburn responde al problema Frege-Geach mediante una noción ampliada de consecuencia lógica, que contempla relaciones de implicación y exclusión entre afirmaciones que no expresan proposiciones en sentido estricto. Ello permite que estas afirmaciones que no describen estados de cosas puedan ser tratadas como si fueran susceptibles de ser verdaderas o falsas y esto, a su vez, permite explicar el hecho de que los hablantes predicen verdad y falsedad de las afirmaciones éticas en su práctica habitual. Como ha señalado Chrisman, el cuasi-realismo puede recurrir a una concepción deflacionaria de la verdad para acomodar los rasgos del discurso ético que motivan la explicación representacionista (Chrisman, 2011: 41).



Ridge mantiene que el expresivismo no necesita asumir que las afirmaciones expresan o bien creencias o bien deseos (Ridge, 2006: 303), sino que hay espacio para una posición híbrida de acuerdo con la cual las afirmaciones éticas expresan, además de estados conativos, estados cognitivos que proporcionan sus condiciones de verdad (Ridge, 2006: 307). Así, Ridge mantiene que cuando un hablante hace una afirmación ética como “Es bueno compartir” expresa, por un lado, un estado conativo a favor de las acciones en general, en la medida en que estas tienen una cierta propiedad, y, por otro lado, la creencia de que la acción de compartir tiene esa propiedad (Ridge, 2006: 313). El componente cognitivo de las afirmaciones éticas permite evitar el problema Frege-Geach, a la vez que el componente conativo da cuenta del comportamiento distintivo de los términos éticos.



Bar-On y Chrisman consideran que el error del emotivismo radica en que explica en términos del contenido semántico la relación que se establece entre las afirmaciones éticas y la motivación de los hablantes. Una teoría del significado del discurso ético debe admitir que las afirmaciones éticas expresan estados conativos para poder explicar la conexión de estas con la motivación, pero no tiene por qué admitir que es la expresión de estados conativos la que confiere significado a estas afirmaciones (Bar-On y Chrisman, 2009: 134). Según dicen, la afirmación ética no solo evidencia un estado mental, sino que también produce una oración que tiene contenido proposicional, y la proposición producida puede ser especificada eliminando las comillas de dicha oración (Bar-On y Chrisman, 2009: 141). De acuerdo con ello, la afirmación “Es bueno compartir” expresa, además de un estado mental, la proposición de que es bueno compartir. Es decir, Bar-On y Chrisman dan cuenta del comportamiento proposicional de las afirmaciones éticas por medio de una noción deflacionaria de proposición.



TERCER PREMIO: Inferencialismo en el discurso ético

Las soluciones de Blackburn, Ridge y Bar-On y Chrisman vienen a subrayar la tensión que se produce entre la tesis de la bifurcación y las continuidades lógico- semánticas que las afirmaciones éticas y las descriptivas presentan. Si se insiste en el contraste entre ambas clases de afirmaciones, como hace el emotivismo, entonces resulta difícil explicar las continuidades. El recurso a nociones deflacionarias de verdad y proposición y la tesis de que las afirmaciones éticas involucran la expresión de estados cognitivos permiten explicar las continuidades, pero diluyen la distinción entre afirmaciones éticas y descriptivas (véase Chrisman, 2011: 43). Así, Bar-On y Chrisman mantienen que no es la expresión de estados conativos la que confiere significado a las afirmaciones éticas y Ridge abandona la dicotomía entre afirmaciones que expresan estados conativos y afirmaciones que describen el mundo para mantener que todas las afirmaciones que pueden ser insertadas en argumentos describen el mundo en alguna medida – incluso si parte de su significado es debido a la expresión de estados conativos–. Otros autores han renunciado de manera más contundente a la tesis de la bifurcación. Schroeder (2008) sostiene que el expresivismo necesita admitir que todas las afirmaciones significativas lo son en virtud de que expresan estados mentales no cognitivos. Por su parte, Price (2013), Chrisman (2008; 2010; 2011) y Frápolli (2016, manuscrito) se apoyan en los problemas que el discurso ético plantea para abogar por una explicación inferencialista del significado de cualquier afirmación.

Este trabajo constituye una defensa de la vía iniciada por Price, Chrisman y Frápolli. La teoría inferencialista del significado permite dar cuenta de los comportamientos que son comunes a las afirmaciones éticas y las afirmaciones descriptivas, y Chrisman ha ofrecido un modo de dibujar la distinción entre ambas clases de afirmaciones en el seno de esa teoría al afirmar que las afirmaciones éticas, a diferencia de las descriptivas, conllevan inferencias que comprometen la acción no lingüística del hablante (Chrisman, 2008: 353). Sin embargo, la propuesta requiere desarrollo. En particular, es necesario explicar en virtud de qué las afirmaciones éticas conllevan ese tipo de inferencias, una vez que se ha renunciado a explicar su significado en términos de expresión de estados conativos. La hipótesis de trabajo es que los términos éticos son predicados de orden superior que toman acciones como argumentos. De ser correcta, esta hipótesis no solo completaría la tesis de Chrisman, sino que, además, explicaría el carácter irreductible de los términos éticos que el argumento de Moore pone de manifiesto.



2. Moore y el argumento de la pregunta abierta

En *Principia Ethica*, Moore presenta el que ha sido denominado argumento “de la pregunta abierta” o argumento “What’s at issue?” (Gibbard, 2012: 9). Según este argumento, “bueno” no puede ser definido en términos de propiedades naturales, puesto que, cualquiera que sea la definición que ofrezcamos en términos de alguna de estas propiedades, es posible preguntar con sentido si es bueno aquello que tiene esa propiedad. Podría ser cierto –dice Moore– que todas las cosas que son buenas tienen, además, alguna propiedad natural, por ejemplo, que todas las cosas buenas son placenteras. Pero, incluso en ese caso, el significado de “placentero” no equivaldría al de “bueno”: el hecho de que comprendemos lo que significa dudar si lo placentero es bueno pone de manifiesto que se trata de dos nociones diferentes (Moore, 1903/2002: 38-39).



Moore se apoya en este argumento para mantener dos tesis. En primer lugar, infiere que “bueno” es indefinible, en el sentido de que no es posible capturar su intensión, ya que no hay ningún conjunto de términos que pueda sustituirlo (Moore, 1903/2002: 31). En consecuencia, dice, no sería posible explicar su significado a quien que no lo conociese ya (Moore, 1903/2002: 29). En segundo lugar, la idea de que es imposible identificar el significado de “bueno” con el de algún término que designe una propiedad natural, unida a la premisa de que “bueno” «no puede significar otra cosa que alguna propiedad de las cosas» (Moore, 1903/2002: 44), lleva a Moore a inferir que “bueno” designa una propiedad no natural (Moore, 1903/2002: 36).

Gibbard ha mantenido que del argumento de Moore no se sigue su tesis ontológica, incluso si se acepta la premisa de que el término “bueno” designa una propiedad. Lo que el argumento de Moore demuestra es la especificidad del concepto de bondad, que no puede ser equiparado con conceptos como el de placer. Pero, según Gibbard, esto es compatible con que ambos conceptos lo sean de la misma propiedad. Dicho de otro modo, Moore no demuestra que “Lo bueno es el placer” es una afirmación falsa; lo que demuestra es que no es una verdad analítica (Gibbard, 2012: 33). Atendiendo a esta precisión de Gibbard, el argumento de Moore plantea un escenario en el que hay que escoger una de las tres opciones siguientes. La conclusión de que los conceptos éticos son irreductibles a conceptos como el de placer puede ser combinada con una explicación representacionalista del significado de las afirmaciones éticas. En ese caso, la primera opción es aceptar que existen propiedades no naturales y que las creencias que involucran los predicados que nombran estas propiedades no naturales son capaces de motivar la acción sin la colaboración de un estado conativo.

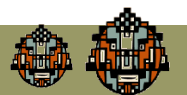
TERCER PREMIO: Inferencialismo en el discurso ético

La segunda opción del representacionista es rechazar el compromiso ontológico anti-naturalista para defender una teoría del error. En tercer lugar, se puede aceptar la conclusión de que los conceptos éticos son irreducibles a conceptos como el de placer y evitar tanto el compromiso con la ontología no naturalista como la teoría del error. En ese caso la única opción es optar por una explicación anti-representacionista de las afirmaciones éticas. De acuerdo con esta explicación, las afirmaciones éticas son significativas, pero no en virtud de que representan estados de cosas y, más concretamente, los predicados éticos son significativos, pero no en virtud de que nombran propiedades. La precisión de Gibbard señala que esta tercera opción es compatible con la tesis de que los predicados éticos designan propiedades. En principio, el anti-representacionista no necesita negar que alguna propiedad caiga bajo la extensión de los conceptos éticos; lo que el anti-representacionista tiene que negar es que estos conceptos sean significativos en virtud de ello. Dicho brevemente, la precisión de Gibbard señala que las posiciones anti-representacionistas como el expresivismo ético no necesitan adquirir compromisos realistas ni anti-realistas respecto de las propiedades éticas. Así pues, conviene reformular el argumento de la pregunta abierta de un modo que evite esos compromisos. Por el momento diremos que el significado de “bueno” no es equivalente al significado de ningún predicado “P”, como “placentero”, puesto que siempre tiene sentido preguntar, dada una afirmación que identifique el significado de “bueno” con el de “P”, si aquello de lo que se predica “P” es bueno.

3. Frege y el argumento de la pregunta abierta. Una explicación no representacionista de las atribuciones de verdad

Si la razón por la que “bueno” es significativo no fuera que representa una propiedad no natural, ¿a qué podría deberse su comportamiento peculiar? En “El pensamiento: una investigación lógica”¹ (1918/1998), Frege propone un argumento a favor del carácter indefinible de la palabra “verdadero” que es análogo al argumento de la pregunta abierta de Moore. Esto proporciona una razón para pensar que la explicación del significado de “verdadero” que Frege ofrece en ese ensayo podría arrojar luz sobre el significado de “bueno”. A continuación se revisará la posición de Frege y se mantendrá, como conclusión de esta revisión, que existen razones para optar por una explicación anti-representacionista del significado de “verdadero”, si bien esa explicación se distancia de la que han propuesto los expresivistas éticos que explican el significado de al menos una clase de expresiones en términos de actitudes o estados mentales.

El argumento fregeano es el siguiente: cualquier intento de definir “verdad” en términos de propiedades o relaciones está abocado al fracaso, puesto que siempre es posible preguntar, respecto de esas propiedades o relaciones que habrían de definir la verdad, si es verdad que se dan. Supongamos, por ejemplo, que definimos la verdad como la correspondencia de una representación con la realidad en un determinado aspecto. En ese caso, para poder decidir si una representación es verdadera tendríamos que preguntarnos si es verdad que se da la correspondencia requerida entre ella y la realidad. El hecho de que sea posible preguntar si es verdad que se da aquello que equivaldría a ser verdadero lleva a Frege a inferir que el contenido de “verdadero” es *sui generis* e indefinible (Frege, 1918/1998: 199), del mismo modo que la inteligibilidad de la pregunta por la bondad de aquello que equivaldría a ser bueno llevaba a Moore a inferir que “bueno” es indefinible.



Así pues, Frege mantiene que no es posible capturar el significado de “verdadero” por medio de una definición. En ese sentido, “verdadero” es para él un término primitivo. Sin embargo, Frege no renuncia a ofrecer una caracterización de lo que considera el uso apropiado del término. En este segundo sentido, sí ofrece una explicación de lo que “verdadero” significa. La palabra “verdadero”, dice, funciona como un predicado; encontramos a la verdad predicada de figuras, representaciones, oraciones y pensamientos (Frege, 1918/1998: 198). Pero, a su juicio, el hecho de que el mismo predicado se aplique tanto a cosas visibles y audibles como a cosas no perceptibles indica que se han producido desplazamientos en el uso del término (Frege, 1918/1998: 198). Cuando el uso es apropiado, dice Frege, aquello respecto de lo cual la verdad entra en consideración es el pensamiento o sentido de una oración (Frege, 1918/1998: 199). De acuerdo con la terminología contemporánea, esto equivale a decir que la verdad es una noción predicativa cuyos argumentos apropiados son proposiciones. Por tanto, se puede afirmar que, en este ensayo², Frege ofrece razones para mantener que el término “verdad” opera como una función de proposiciones. Esta interpretación se ve apoyada por su idea de que “verdadero” no atribuye una propiedad como las que denominamos con palabras como “rojo” o “amargo” (Frege, 1918/1998: 200).

2 En otros lugares, Frege defiende que la verdad es un objeto y que el término “verdad” es el nombre de ese objeto (véanse Frege, 1891/1984; 1892/2005). Sin embargo, la explicación del significado de “verdadero” que propone en “El pensamiento” apoya la tesis de que “verdadero” es una función de proposiciones y, como se argumenta más abajo, constituye una razón para atribuir a Frege una concepción no representacionista de la verdad.

1 En adelante, “El pensamiento”.

TERCER PREMIO: Inferencialismo en el discurso ético



¿Cuál es, entonces, el significado de la palabra “verdad”? La respuesta que Frege proporciona consiste en una especificación de la función que desempeñan las atribuciones de verdad. Según dice, la oración “Es verdad que huelo aroma de violetas” tiene el mismo contenido que la oración “Huelo aroma de violetas”: el predicado de verdad no añade nada al pensamiento expresado por la segunda oración. Sin embargo, Frege no considera que la función de la atribución de verdad sea trivial o redundante. Al contrario, afirma que obtenemos «un gran resultado» cuando concluimos que es verdadero algo que habíamos conjeturado (Frege, 1918/1998: 201). Para comprender la concepción fregeana de la noción de verdad hay que atender a la distinción que Frege establece entre las acciones de pensar, juzgar y aseverar: pensar consiste en captar un pensamiento, sin proponerlo como verdadero; juzgar consiste en reconocer la verdad del pensamiento; por último, la acción de aseverar consiste en manifestar el juicio o reconocimiento de la verdad del pensamiento (Frege, 1918/1998: 202). De acuerdo con esta distinción, hacer una aserción consiste ya en proponer un pensamiento como verdadero y el hablante puede hacer esto sin usar la palabra “verdad”. Es decir, la afirmación de la verdad de un pensamiento es una acción que no requiere que la palabra “verdad” sea usada (Frege, 1918/1998: 203). La consecuencia parece ser que la palabra “verdad” se utiliza para hacer explícita la acción que el hablante realiza por medio de la aserción de un pensamiento. Así pues, aunque la palabra “verdad” no contribuye al pensamiento expresado, la función que la atribución de verdad desempeña, que es equivalente a la función que desempeña la aserción, es irreductible (véase Frápolli, 2013: 5).

En definitiva, la explicación de Frege sugiere que “verdadero” es una función de proposiciones que sirve para hacer explícita la acción que el hablante realiza con las proposiciones que saturan la función, a saber, la aseveración de esas proposiciones. Esta explicación permite comprender el porqué del carácter desconcertante del predicado de verdad que se pone de manifiesto en el argumento fregeano de la pregunta abierta. Si la explicación de Frege es correcta, la expresión “verdadero” es una expresión funcional de orden superior (véanse Frápolli, 2013: 8; Frápolli y Villanueva, 2013: 595), mientras que las expresiones que habrían de sustituirla son expresiones funcionales de primer orden.

La distinción que aquí se establece entre funciones de primer orden y funciones de orden superior se distancia de la que establece el propio Frege. Según Frege, las funciones de primer orden son aquellas cuyos argumentos son objetos o entidades completas, mientras que las expresiones de segundo orden toman otras funciones como argumentos (Frege, 1891/1984: 43). Es decir, los argumentos de las funciones de segundo orden son, de

acuerdo con la distinción de Frege, entidades incompletas. En cambio, las proposiciones son entidades que no necesitan ser completadas por ninguna variable y, así, las funciones de proposiciones serían funciones de primer orden si se atendiera a la distinción fregeana. De acuerdo con la distinción que aquí se va a emplear, las funciones de orden superior son aquellas que toman como argumentos predicables y oraciones –desde el punto de vista del lenguaje–, o conceptos y proposiciones –desde el punto de vista del contenido– (Frápolli y Villanueva, 2013: 595). Por tanto, diremos que las funciones de proposiciones son de orden superior.

Uno de los rasgos que permiten identificar a las funciones de orden superior y distinguirlas de otras de primer orden es que solo las primeras pueden tomar oraciones como argumentos (Frápolli y Villanueva, 2012: 475). Frege oponía “verdadero” a expresiones como “rojo” y “amargo” y, de hecho, en oraciones como “Es verdad que huelo aroma de violetas”, la expresión “es verdad” no puede ser sustituida por expresiones como “es rojo” o “es amargo” sin que la gramaticalidad de la oración se vea afectada. Este hecho constituye una prueba de que los términos “verdadero”, por un lado, y “rojo” y “amargo”, por otro, pertenecen a categorías diferentes (véase Frápolli y Villanueva, 2012: 476). De entre esos tres predicados, solo “verdadero” es de orden superior. “Rojo” y “amargo”, en cambio, son predicados de primer orden.

Decíamos más arriba que, si la explicación fregeana de las atribuciones de verdad es correcta, el predicado de verdad es de orden superior, mientras que las expresiones que habrían de proporcionar su significado son predicados de primer orden. La expresión “correspondiente con la realidad”, entendida en un sentido puramente descriptivo, es una función de primer orden que toma representaciones como argumentos. Ello explica por qué la expresión “verdadero” no puede ser sustituida por una expresión como “correspondiente con la realidad”: se trata de expresiones que desempeñan diferentes funciones lógicas. La lección que arroja el argumento de la pregunta abierta de Frege es, entonces, que un predicado de orden superior no es equiparable a ningún conjunto de predicados de primer orden.

Otra de las consecuencias que se desprenden del planteamiento de Frege es que la teoría representacionalista del significado no explica adecuadamente el significado de palabras como “verdadero”. Frege mantiene que el término “verdadero” no contribuye a la proposición expresada por la afirmación en la que aparece –el contenido de “Es verdad que huelo aroma de violetas” es el mismo que el de “Huelo aroma de violetas”–. Atribuir verdad a una proposición consiste, de acuerdo con Frege, en hacer explícita la aseveración de la misma por parte del hablante, no en describir la proposición. En ese sentido, su función es expresiva.

TERCER PREMIO: Inferencialismo en el discurso ético

El sentido en el que cabe decir que para Frege el término “verdadero” tiene una función expresiva se distancia del que emplean expresivistas como Ayer y Stevenson. Las posiciones expresivistas de acuerdo con las cuales hay una clase de afirmaciones que son significativas en virtud del estado mental que los hablantes expresan por medio de ellas proporcionan un tratamiento no representacionista del término “verdadero”. Pero la tesis fregeana de que la función de “verdadero” no es descriptiva no le compromete con todas las tesis que mantienen estos expresivismos. En pocas palabras, el paso del anti-representacionismo al tipo de expresivismo mencionado no es necesario.



Frápolti y Villanueva han distinguido cuatro tesis que suelen ir asociadas a las posiciones expresivistas y han mantenido que basta con que una posición acepte dos de ellas para que pueda ser considerada expresivista (Frápolti y Villanueva, 2012: 471).

Las tesis son las siguientes: (i) hay expresiones del lenguaje natural que son funciones de proposiciones y estas son expresiones que no afectan a las condiciones de verdad; (ii) esas expresiones no se utilizan para describir cómo es el mundo; (iii) las afirmaciones que contienen alguna de estas expresiones carecen de condiciones de verdad; y (iv) esas expresiones son usadas para expresar una actitud hacia un contenido. Las tesis (i), (ii), (iii) y (iv) son lógicamente independientes (Frápolti y Villanueva, 2012: 471), de modo que sería posible mantener (i) y (ii) sin mantener (iii) y (iv). Esta es la opción por la que ellos mismos abogan, a la que denominan “expresivismo mínimo” (Frápolti y Villanueva, 2012: 474). El expresivismo mínimo es la conjunción de (i), (ii) y una quinta tesis que es consecuencia de (ii): (v) los predicados de orden superior no modifican las condiciones de verdad de las expresiones que caen bajo su alcance (Frápolti y Villanueva, 2012: 478). El compromiso con (iv) no es necesario, porque el significado de las funciones de proposiciones puede ser explicado de manera inferencialista: si sabemos qué se sigue de la aplicación de un concepto y de qué se sigue la aplicación de un concepto, entonces sabemos cuál es el significado del concepto; no necesitamos nada más (Frápolti y Villanueva, 2012: 481). Así, el expresivismo mínimo hace posible una explicación no representacionista de las atribuciones de verdad que se distancia de la que ha defendido el expresivismo clásico para las atribuciones de bondad. Podemos mantener que las atribuciones de verdad y el término “verdad” son significativos porque están insertos en relaciones inferenciales. Quizá es cierto que por medio de las atribuciones de verdad el hablante expresa una actitud o un estado mental, pero no es ese estado mental el que confiere significado a la afirmación, ni a los términos que esta involucra.

Brandom (2000/2002) ha mantenido que, de hecho, Frege es partidario de la concepción inferencialista del significado. La *Conceptografía*, dice, tiene como objetivo la explicación de lo que Frege llama “contenido conceptual”, y la calificación de “conceptual” se interpreta en términos inferenciales: dos afirmaciones tienen el mismo contenido conceptual si, y solo si, tienen la misma función inferencial (Brandom, 2000/2002: 63). El propio Frege afirma que dos juicios tienen el mismo contenido cuando todas las consecuencias que se derivan de la combinación de uno de ellos con otros juicios se derivan también de la combinación del otro con los mismos juicios (Frege, 1879/1972: 14-15). Es esa parte del contenido que es común a ambos juicios la que Frege encuentra relevante para la conceptografía, mientras que aquellos recursos lingüísticos que no afectan a las consecuencias de los juicios no son formalizados en su lenguaje. Según dice, estos recursos sirven para facilitar al oyente la comprensión del contenido, para ponerle sobre la pista correcta, pero no contribuyen al contenido mismo (Frege, 1879/1972: 15). Si la interpretación propuesta por Brandom es correcta³, entonces la posición de Frege es que “Huelo aroma de violetas” y “Es verdad que huelo aroma de violetas” adquieren sus contenidos en virtud de sus respectivas relaciones inferenciales, de tal modo que ambas afirmaciones tienen el mismo contenido si, y solo si, sus respectivas relaciones inferenciales coinciden. Según esto, la tesis de Frege de que ambas afirmaciones tienen el mismo contenido –la tesis de que la palabra “verdad” no modifica el contenido de la oración “Huelo aroma de violetas”– equivale a la tesis de que ambas afirmaciones conllevan las mismas relaciones inferenciales. Así, hay que concluir que aquello que se sigue de la atribución –explícita– de verdad es lo mismo que se sigue de la mera afirmación del contenido; y que aquello de lo que la atribución –explícita– de verdad se sigue es lo mismo de lo que se sigue la mera afirmación. Con ello, la interpretación inferencialista de la explicación que Frege ofrece en “El pensamiento” proporciona un trasfondo teórico a la tesis fregeana de que atribuir verdad a un pensamiento no es otra cosa que afirmarlo.

Para recapitular, la caracterización del uso apropiado del término “verdadero” que Frege hace en “El pensamiento” proporciona razones para individualizar el concepto de verdad atendiendo a las relaciones inferenciales en que está inmerso y permite mantener, coherentemente con Brandom y con el expresivismo mínimo, que la función que Frege atribuye al término “verdadero” es expresiva en el sentido de que el término expresa o explicita algo que está implícito. No se quiere decir con ello que

3 La interpretación inferencialista de la discusión fregeana acerca de las atribuciones de verdad se ve apoyada por algo que Frege dice en “El pensamiento”, a saber, que el significado de la palabra “verdad” se despliega en las «leyes del ser verdad» (Frege, 1918/1998: 197), de las que se siguen «prescripciones para afirmar, pensar, juzgar e inferir» (Frege, 1918/1998: 196).

TERCER PREMIO: Inferencialismo en el discurso ético

Frege atribuya al término la función de expresar estados mentales y, sobre todo, no se quiere decir que sea la expresión de estados mentales la que confiere al término su significado; nada en el planteamiento de Frege hace pensar que los estados mentales de los hablantes son relevantes cuando se trata de individuar el significado. Por el contrario, la posición de Frege es compatible con el expresivismo mínimo. En primer lugar, Frege mantiene que la verdad es un predicado cuyos argumentos apropiados son proposiciones y que la predicación de verdad no modifica la proposición sobre la que actúa. Con ello admite una versión de la tesis (i), de acuerdo con la cual hay expresiones que actúan como funciones de proposiciones y que no contribuyen a la proposición expresada por la afirmación en la que aparecen. En segundo lugar, mantiene que afirmar que algo es verdadero no consiste en atribuir una propiedad a ese algo, sino en manifestar la aseveración de una proposición, con lo que admite, para el caso de “verdadero”, la tesis (ii), según la cual las expresiones que actúan como funciones de proposiciones no se usan para describir estados de cosas. Por otro lado, Frege rechaza (iii) y no se compromete con (iv). De acuerdo con el método de individuación del contenido que propone su conceptografía, que las afirmaciones “Es verdad que huelo aroma de violetas” y “Huelo aroma de violetas” tienen el mismo contenido significa que todas las proposiciones que se siguen de “Es verdad que huelo aroma de violetas” se siguen también de “Huelo aroma de violetas” y, viceversa, que todas las proposiciones que se siguen de “Huelo aroma de violetas” se siguen también de “Es verdad que huelo aroma de violetas”. El rechazo de (iii) es una consecuencia de la idea de que algo se sigue de “Es verdad que huelo aroma de violetas”: si algo puede seguirse de esa afirmación, ha de ser que esta expresa una proposición y, por tanto, el operador de verdad no cancela la proposición sobre la que actúa (véase Frápolli y Villanueva, 2012: 477-478). En cuanto a (iv), Frege concibe “verdadero” como un término cuya función no es representar, pero ello no le compromete con la idea de que el término expresa un estado mental del hablante.

4. Una explicación no representacionista de las atribuciones de bondad

El argumento fregeano de la pregunta abierta muestra que el significado de “verdadero” no puede ser reducido al de ningún conjunto de términos que sirvan para atribuir propiedades naturales. La discusión previa ha estado encaminada a establecer que la razón de esta irreductibilidad es que “verdadero” es un predicado de orden superior, mientras que los términos que atribuyen propiedades naturales son predicados de primer orden. Cuando el argumento de la pregunta abierta de Frege se examina a la luz de su explicación del papel que desempeñan las atribuciones de verdad, la conclusión que se sigue del argumento es que un predicado de orden superior no es reducible a ningún conjunto de predicados de primer orden. La función que realizan los predicados de orden superior no puede ser realizada por ningún

conjunto de predicados que atribuyan propiedades naturales. Ello proporciona un argumento a favor de la explicación anti-representacionista del significado del predicado de verdad.



Podemos retomar ahora la cuestión del significado del discurso ético. El paralelismo que se establece entre los argumentos de la pregunta abierta de Moore y de Frege sugiere que los predicados éticos, al igual que el predicado de verdad, son funciones de orden superior. Ello explicaría por qué ningún conjunto de términos cuya función es atribuir propiedades naturales es capaz de capturar su significado. Por otro lado, la explicación de las atribuciones de verdad que Frege propone como solución a su argumento de la pregunta abierta, unida al paralelismo que se establece entre este argumento y el de Moore, sugiere que la explicación adecuada del significado de los predicados éticos podría ser no representacionista.

También Gibbard se ha apoyado en el argumento de la pregunta abierta de Moore para establecer una distinción entre dos tipos de conceptos por medio de los cuales los hablantes realizarían diferentes acciones y para argumentar a favor de una explicación no representacionista de los conceptos de uno de esos tipos. En su caso, la distinción se establece entre lo que él llama «conceptos naturalistas» y «conceptos normativos». Lo que el argumento de Moore pondría de manifiesto es que los conceptos normativos no pueden ser explicados en términos puramente naturalistas (véase Gibbard, 2012: 33, 92). Según Gibbard, la razón de esta irreductibilidad es que los conceptos normativos involucran el concepto de deber, que no se sigue analíticamente del plano naturalista de las cosas (Gibbard, 2012: 11). Por ello, dice, el significado de “deber” requiere una explicación «expresivista». La idea es que, dado que no es posible especificar el significado de “deber” por medio de términos naturalistas o no normativos, es preciso ofrecer una explicación indirecta de ese significado, que indique qué estado mental equivale a significar “deber” mediante una palabra (Gibbard, 2012: 223). Es decir, Gibbard coincide con los expresivistas clásicos en caracterizar la función expresiva de las palabras no descriptivas en términos de estados mentales de los hablantes. En particular, Gibbard mantiene que los pensamientos acerca de deberes o pensamientos normativos son conceptualmente equivalentes a pensamientos de planificación (Gibbard, 2012: 194); el acto que los hablantes realizan por medio de las afirmaciones normativas consiste en la expresión de un compromiso para la acción, y no en una descripción del mundo. Si alguien cree que debe hacer x, su creencia es conceptualmente equivalente al plan de hacer x; si alguien cree que otra persona debe hacer x, su creencia es conceptualmente equivalente al plan de hacer x en caso de estar en el lugar de esa otra persona (véase Gibbard, 2012: 170).

TERCER PREMIO: Inferencialismo en el discurso ético

Así, tener una creencia normativa y rechazar el correspondiente pensamiento de planificación –o viceversa– es cometer una incoherencia conceptual (Gibbard, 2012: 194), el mismo tipo de incoherencia que cometería quien creyera que John es soltero y está casado. Aún más, planear hacer algo y *ceteris paribus* no hacerlo supone cometer una incoherencia de este tipo (Gibbard, 2012: 206). Esto permite explicar por qué los hablantes presentan la motivación de conducirse de acuerdo con sus afirmaciones éticas. Si los pensamientos éticos equivalen a pensamientos de planificación, entonces un hablante que hace una afirmación ética a pesar de que no planea actuar de acuerdo con ella, o bien es un hablante insincero que no hace el juicio que expresa, o bien es un hablante incompetente.

La propuesta de entender las afirmaciones normativas como afirmaciones conceptualmente equivalentes a planes para la acción proporciona una vía de desarrollo de la tesis de que los predicados éticos son de orden superior, ya que sugiere una hipótesis acerca de qué tipo de argumento satura estos predicados. A partir de la analogía entre los argumentos de la pregunta abierta de Frege y de Moore se ha concluido que, al igual que “verdadero”, “bueno” actúa como una función de orden superior cuando se emplea en un sentido ético. A pesar de que encontramos usos éticos en los que la bondad parece ser un predicado de primer orden, como en “John es un niño bueno”, la conclusión se ve confirmada por el hecho de que la expresión “es bueno” puede ser saturada con oraciones. Así ocurre, por ejemplo, en afirmaciones como “Es bueno que John comparta el juguete con su hermano”.

www.solofici.org

Pero, a diferencia del predicado de verdad, el predicado de bondad podría no ser una función de proposiciones de modo fundamental. La razón es que también son posibles las afirmaciones éticas en las que la expresión funcional “Es bueno” va seguida de expresiones suboracionales, no susceptibles de expresar proposiciones. Es el caso de afirmaciones como “Es bueno compartir”. Si el predicado de bondad es de orden superior y no es una función de proposiciones, entonces ha de ser una función de conceptos o predicados de primer orden. Es este consecuente el que encuentra apoyo en el planteamiento de Gibbard. Según Gibbard, «las cuestiones morales son cuestiones acerca de cómo sentirse respecto de las cosas que la gente hace o podría hacer» (Gibbard, 2003/2009: xi). Esto es, las afirmaciones éticas constituyen una toma de posición del hablante ante cursos de acción posibles. Si esto es cierto, entonces los argumentos apropiados del predicado de bondad son acciones. Los usos éticos del término “bueno” en los que este parece actuar como un predicado de primer orden pueden ser acomodados dentro de esta explicación que identifica las afirmaciones éticas con afirmaciones acerca de planes para la acción. Que “John es un niño bueno” significa que John es un niño que se comporta bien, o que tiene los sentimientos adecuados respecto de determinados cursos de acción

posibles. Esto es, que John es un niño cuyas acciones o predisposiciones para la acción son buenas. Así, estos usos éticos en los que “bueno” actúa como un predicado de primer orden pueden ser explicados como el resultado de un desplazamiento a partir de un uso más fundamental, como Frege hiciera para el caso de la verdad.



La intención no es afirmar que todos aquellos conceptos que implican deberes son funciones cuyos argumentos son acciones. Gibbard llama “conceptos normativos” a los conceptos que no son reducibles a conceptos descriptivos o naturalistas, de modo que los conceptos éticos son un subconjunto de los conceptos normativos. De acuerdo con él, todas las afirmaciones que involucran conceptos normativos implican deberes y, en ese sentido, compromisos prácticos. Pero de esto no se sigue que todas las afirmaciones que involucran conceptos normativos constituyan una toma de posición del hablante ante una acción. El concepto de verdad proporciona un contraejemplo. El argumento de la pregunta abierta de Frege establece que el de verdad es uno de esos conceptos que Gibbard considera normativos y, sin embargo, los argumentos del predicado de verdad son proposiciones. Por medio del predicado de verdad un hablante hace explícita su toma de posición ante un contenido juzgable, no ante una acción. ¿Conllevan las adscripciones de verdad deberes o compromisos para la acción? La respuesta que se desprende del discurso de Frege es afirmativa: si la adscripción de verdad a una proposición consiste en hacer explícito el acto de afirmar la proposición, entonces la adscripción de verdad a una proposición implica el deber de asumir todas aquellas proposiciones que se siguen inferencialmente de la proposición que se juzga verdadera, así como el deber de rechazar todas aquellas proposiciones que son incompatibles con la proposición que se juzga verdadera.

La hipótesis de que el predicado de bondad es una función de orden superior que toma acciones como argumentos permite dar cuenta de la tendencia de los hablantes a actuar de acuerdo con los juicios éticos que realizan de un modo que se distancia de la que ha sido la explicación expresivista tradicional. Los expresivistas éticos han explicado esa conexión de las palabras con la acción a través de la tesis de que las afirmaciones éticas adquieren su significado en virtud de la expresión de estados conativos. La idea es que si las afirmaciones éticas consisten en la expresión de estados conativos, entonces no es sorprendente que los hablantes que hacen afirmaciones éticas tengan los estados mentales que otorgan significado a esas afirmaciones. Gibbard se suma a esta explicación. Pero, una vez más, no es necesario apelar a estados mentales para explicar el significado de las afirmaciones que no representan estados de cosas. Lo distintivo de las afirmaciones normativas es su vinculación conceptual con la acción, en palabras del propio Gibbard (2012: 227), y la vinculación conceptual entre la acción y las palabras

TERCER PREMIO: Inferencialismo en el discurso ético

del hablante puede ser capturada inferencialmente. Así lo hace Chrisman, que establece un contraste entre afirmaciones éticas y afirmaciones descriptivas atendiendo, no a los tipos de estados mentales que sendas clases de afirmaciones expresarían –estados conativos frente a estados cognitivos, respectivamente–, sino a los tipos de relaciones inferenciales que sendas clases de afirmaciones harían explícitos (Chrisman, 2011: 50). La especificidad de los conceptos éticos se debe a que estos tienen un papel inferencial práctico, mientras que los conceptos descriptivos están insertos en relaciones inferenciales de carácter teórico (Chrisman, 2008: 353). Así, las afirmaciones normativas adquieren su significado por la misma vía que cualquier otra afirmación, a saber, en virtud de sus relaciones inferenciales, pero solo las afirmaciones normativas tienen una conexión inferencial con la acción (Chrisman, 2008: 353). La hipótesis de que los predicados éticos son funciones de orden superior cuyos argumentos son acciones permite completar esta propuesta de Chrisman, ya que explica a qué es debida la conexión inferencial de las afirmaciones éticas con la práctica no lingüística de los hablantes. Los términos como “bueno” son recursos lingüísticos por medio de los cuales los hablantes toman posición ante cursos de acción posibles. En ese sentido, las afirmaciones éticas expresan los planes de los hablantes.

La propuesta expresivista que resulta de la reformulación en términos inferencialistas de la tesis de Gibbard es una forma de no naturalismo que no conlleva una ontología anti-naturalista. En palabras de Gibbard, no es que haya propiedades no naturales, es que hay conceptos no naturalistas (Gibbard, 2012: 229). Los conceptos éticos surgen como consecuencia del hecho de que los humanos somos animales que hacen planes, capaces de pensar acerca de qué hacer ahora o en contingencias futuras (Gibbard, 2003/2009: xi). Y hablar acerca de qué planeamos hacer es tan legítimo como hablar acerca de cómo son las cosas (Gibbard, 2012: 224). El lenguaje tiene otras funciones además de la de describir cómo es el mundo.

5. Funciones de predicados y el problema Frege-Geach

Hasta aquí se ha argumentado que “bueno” es un predicado de orden superior que toma acciones como argumentos y que ello sugiere que la explicación de las afirmaciones éticas debe ser no representacionalista. La propuesta satisface dos de las motivaciones tradicionales del expresivismo ético: por un lado, explica el carácter aparentemente no naturalista de los conceptos éticos; por otro, da cuenta de la motivación de los hablantes a actuar de modo coherente con sus afirmaciones éticas. Pero la propuesta deber satisfacer un tercer requisito para constituir una explicación viable del significado de las afirmaciones éticas, a saber, debe ser capaz de hacer frente al problema Frege-Geach. Como se ha dicho más arriba, el argumento Frege-Geach pone de manifiesto que las oraciones éticas pueden

actuar como premisas y conclusiones en argumentos válidos y que pueden caer bajo el alcance de operadores veritativo-funcionales, como el condicional y otras conectivas lógicas. En consecuencia, una teoría del significado de las afirmaciones éticas debe poder atribuir a estas contenido proposicional.

A primera vista podría parecer que la hipótesis de que “bueno” actúa como una función de predicados no permite dar cuenta del contenido proposicional de las afirmaciones éticas y, por tanto, que el argumento de Geach nos obliga a abandonar esta hipótesis. Si “bueno” es una función de predicados, una afirmación como “Es bueno compartir” estaría compuesta por dos expresiones predicativas que, en principio, no bastan para componer un contenido proposicional. La primera de ellas es la expresión “es bueno”, que no se usaría para describir cómo es el mundo y que, por ello, no contribuiría a las condiciones de verdad de la afirmación en la que aparece. La segunda expresión es “compartir”, que no es susceptible de expresar una proposición.

La irrelevancia para las condiciones de verdad de los predicados de orden superior no plantea este problema a las afirmaciones que involucran el predicado “verdadero”, porque “verdadero” opera sobre contenidos proposicionales. Incluso si “es verdad” no contribuye a la proposición expresada por la afirmación “Es verdad que huelo aroma de violetas”, esa afirmación expresa una proposición, a saber, que huelo aroma de violetas. Una posible vía de solución del problema que surge en el caso de las afirmaciones éticas consiste en reformular las afirmaciones como “Es bueno compartir” de modo que su estructura coincida con la de las adscripciones de verdad. La afirmación “Es bueno compartir” sería una versión abreviada de la afirmación “Es bueno que las personas compartan”. Así, incluso si “es bueno” no contribuye a las condiciones de verdad, la afirmación ética expresa una proposición, a saber, que las personas comparten.



¿Significa esto que “bueno” es una función de proposiciones, como lo es “verdadero”? Arriba se ha aceptado que “verdadero” y “bueno” no pueden ser sustituidos por términos que atribuyan propiedades naturales, porque son predicados de orden superior que sirven para indicar la posición que el hablante adopta ante algún tipo de entidad. Es decir, son términos cuya función es normativa. “Verdadero” es un término normativo que indica la posición que el hablante adopta ante un contenido juzgable o proposición y, coherentemente, “verdadero” es una expresión funcional que toma proposiciones como argumentos. Si “bueno” es también una expresión funcional que se satura con proposiciones de modo fundamental, cabría esperar que sirviera al hablante para indicar su toma de posición ante una proposición. Sin embargo, no resulta fácil imaginar en qué sentido cabe atribuir bondad a una proposición, ni qué compromisos inferenciales se seguirían de esa atribución. La opción más intuitiva consiste



TERCER PREMIO: Inferencialismo en el discurso ético

en suponer que el hablante se compromete con el estado de cosas que la proposición representa. De acuerdo con esto, la solución al problema Frege-Geach consistiría en suponer, primero, que dado que la afirmación “Es bueno compartir” puede formar parte de un argumento válido, esa afirmación es equivalente a la de “Es bueno que las personas compartan” y, segundo, que por medio de cualquiera de esas afirmaciones los hablantes indican su adopción de algún tipo de compromiso con el estado de cosas representado por la afirmación en su forma no abreviada. En el ejemplo considerado el hablante adoptaría algún tipo de compromiso con un estado de cosas en el que las personas comparten. ¿En qué consistiría ese compromiso? Dado que los hablantes presentan de manera regular la motivación a actuar de modo coherente con sus juicios éticos, hay que suponer que el compromiso que el hablante adquiere por medio de la adscripción de bondad es el compromiso de promover el estado de cosas representado por la proposición que saturaría el predicado de bondad.

www.solofici.org

Si “Las personas comparten” representa un estado de cosas, entonces la solución ensayada podría dar cuenta del contenido proposicional de las oraciones éticas, así como del carácter no naturalista de los conceptos morales y de la vinculación que se establece entre los juicios éticos y la conducta no lingüística de los hablantes que hacen esos juicios. Sin embargo, la solución requiere que ejerzamos una cierta violencia sobre el modo en que los hablantes emplean los términos éticos en su práctica habitual y, con todo, no logra rechazar la intuición de que por medio de las afirmaciones éticas los hablantes se posicionan ante determinados cursos de acción, ya que expresar el compromiso de promover un estado de cosas en el que todas las personas comparten parece equivalente a expresar el compromiso de compartir en caso de estar en el lugar de cualquier persona. Frápolli (2016, manuscrito) ha mantenido que, incluso si una afirmación general como “Es bueno compartir” se reformula de modo que “Es bueno” sea saturado con una cláusula-que, la afirmación ética no representa una situación y, en ese sentido, no expresa una proposición. La reformulación que ella considera es “Es malo que tú tortures animales”. Según dice, “tú” debería referir a un individuo para que la cláusula-que representara una situación, pero el uso de “tú” no es referencial en la afirmación ética que resulta de la reformulación de una afirmación general; “tú” funciona como un cuantificador universal o herramienta de generalización.



El argumento de Frápolli permite explicar por qué resulta contraintuitivo afirmar que la bondad se predica de proposiciones, a pesar de la forma oracional de la cláusula-que que satura la expresión “es bueno”. En el

ejemplo que se ha venido discutiendo, la expresión “las personas”, que se añade a la afirmación ética general “Es bueno compartir” con el objetivo de construir una proposición, opera como una herramienta de cuantificación y no como un término referencial que seleccione una extensión. Así, incluso cuando la forma de las adscripciones de bondad es tal que la expresión “es bueno” es saturada con una oración, la bondad no se predica de proposiciones. Los argumentos apropiados del predicado de bondad son cursos generales de acción (Frápolli, 2016, manuscrito) y, por tanto, la reformulación de las afirmaciones éticas no permite resolver el problema Frege-Geach (Frápolli, 2016, manuscrito).

6. El anti-representacionalismo como tesis global

Llegados a este punto, se podría pensar que hay que renunciar al proyecto de capturar inferencialmente el significado de los términos éticos y optar por el tipo de anti-representacionalismo que explica su significado en términos de expresión de estados conativos. Si así se hiciera, sería posible dar cuenta del contenido proposicional de las afirmaciones éticas, bien por medio de nociones deflacionarias de proposición y de verdad, o bien por medio de la postulación de un componente descriptivo que estaría involucrado en las afirmaciones éticas. Pero todavía cabe dar un paso en la dirección opuesta: en lugar de renunciar al inferencialismo, podemos llevar hasta sus últimas consecuencias la estrategia inferencialista de individuación del significado, de acuerdo con la cual la unidad mínima de significado es la proposición o contenido juzgable. La estrategia ya se ha esbozado en el apartado 3 de la mano de Brandom y Frege. En la *Conceptografía*, Frege adopta un método de individuación del significado que pregunta en primer lugar por el contenido de los juicios, no por el contenido de las palabras que forman parte de esos juicios (Frege, 1879/1972: 14-15). Por su parte, Brandom considera que son los resultados de los actos de habla asertivos los que adquieren significado en primer lugar, en virtud de que actúan como premisas y conclusiones en inferencias (véase Brandom, 2000/2002: 58). Esta estrategia de Frege y Brandom consistente en otorgar prioridad semántica al acto de habla es una consecuencia de la tesis –que ambos mantienen– según la cual son las relaciones inferenciales las que permiten individuar el significado, ya que las palabras que forman parte de los juicios no son entidades susceptibles de establecer relaciones inferenciales por sí mismas. De la preferencia de una palabra aislada no se sigue nada y, por tanto, esta no puede actuar como premisa de un argumento. Esto no implica que haya que renunciar a la tesis de que los conceptos están inmersos en relaciones inferenciales que proporcionan su significado. Los conceptos están involucrados en relaciones inferenciales, pero lo están por mediación de las proposiciones en las que aparecen y, en consecuencia, el inferencialista ha de individuarlos de manera derivada.

TERCER PREMIO: Inferencialismo en el discurso ético

Frápolti y Villanueva (2015) denominan «modelo orgánico» a esta estrategia de individuación del contenido proposicional compartida por Frege y Brandom, y le oponen una segunda estrategia de individuación, a la que denominan «modelo *building-block*». El modelo *building-block* individúa el contenido de las proposiciones por medio de un proceso de construcción del significado que opera de abajo a arriba, esto es, que pregunta primero por el significado de los componentes suboracionales de una oración para obtener a partir de ellos el significado de la oración que estos componen (Frápolti y Villanueva, 2015: 2). Frente a él, el modelo orgánico toma como punto de partida aquellas entidades que pueden ser usadas para adquirir compromisos inferenciales, esto es, aquellas entidades que pueden actuar como premisas o conclusiones en inferencias (Frápolti y Villanueva, 2015: 3). A diferencia del modelo *building-block*, el modelo orgánico de la individuación del contenido proposicional no requiere que las partes de la proposición respondan a una estructura proposicional preestablecida (Frápolti y Villanueva, 2015: 3), lo cual no quiere decir que las proposiciones no sean analizables. El modelo orgánico es compatible con el análisis proposicional, pero entiende que el análisis ha de ser realizado *a posteriori*, una vez que el contenido proposicional ha sido individuado (Frápolti y Villanueva, 2015: 4-5).

Una consecuencia que se desprende del modelo orgánico de la individuación es que si una entidad tiene el comportamiento propio de una proposición, entonces se puede afirmar que esa entidad es una proposición (Frápolti, 2016, manuscrito) –aún más, se puede decir qué proposición es–, con independencia de cuáles sean las partes que la componen. Así, el inferencialismo ofrece una solución al problema que el comportamiento proposicional de las afirmaciones éticas ha planteado a la hipótesis de que “bueno” es una función de predicados. No necesitamos encontrar componentes subproposicionales que permitan construir un contenido con una cierta estructura para explicar por qué “Es bueno compartir” puede formar parte de un argumento válido como el que propone el argumento de Geach. La estrategia puede ser la inversa. Podemos tomar el hecho de que “Es bueno compartir” puede actuar como premisa de ese argumento –así como el hecho de que puede actuar como antecedente de un condicional– como evidencia de que esa afirmación tiene contenido proposicional (Frápolti, 2016, manuscrito) y, en su caso, inferir el significado de los términos éticos a partir de su contribución a las afirmaciones en las que aparecen. En las secciones previas se ha perseguido una explicación expresivista local del discurso ético, es decir, una explicación que permitiera conservar el representacionalismo como explicación del significado de otras clases de expresiones. Sin embargo, el representacionalismo conlleva un modelo *building-block* de individuación de las proposiciones (véase Frápolti y Villanueva, 2013: 584), que ha resultado ser incapaz de dar cuenta del comportamiento proposicional de las afirmaciones éticas.

Siguiendo a Frápolti (2016, manuscrito), me apoyo en esa incapacidad del representacionalismo para rechazar que este constituya una explicación adecuada del significado de alguna clase de expresión. Hay expresiones que representan, pero ninguna adquiere su significado en virtud de ello. La misma posición ha sido defendida por Price, que ha mantenido que el contraste que el expresivismo establece entre dos tipos de afirmaciones que adquieren su significado por distintas vías conserva presupuestos representacionalistas que los expresivistas no necesitan asumir, ya que el expresivismo es compatible con la noción de proposición que proporciona el inferencialismo de Brandom (Price, 2013: 31). También Price mantiene que hay que renunciar al representacionalismo en todo caso, para hacer del expresivismo una tesis global (Price, 2013: 30).

El modelo de proposición propio del inferencialismo permite ofrecer una nueva respuesta a la que ha sido la pregunta fundadora del expresivismo ético. Los primeros expresivistas evitaron el problema que el discurso ético planteaba a la teoría representacionalista del significado mediante un cambio en la pregunta: su estrategia consistió en preguntar, no cuál es la definición de “bueno”, sino en qué consiste juzgar que algo es bueno (Gibbard, 2003/2009: 6). Es decir, los expresivistas éticos comenzaron por dar prioridad al acto de habla frente al término ético. El inferencialismo conserva la idea de que hay que preguntar por el significado del acto de habla en primer lugar, pero introduce una novedad en la respuesta a esta pregunta. A pesar de que algunas de nuestras afirmaciones describen y otras no, hay un sentido en el que lo que los hablantes hacen por medio de una afirmación ética es lo mismo que hacen por medio de cualquier otra afirmación, a saber, realizan un movimiento en el juego inferencial de dar y pedir razones. Price ha afirmado que el expresivismo debe seguir a Brandom en la idea de que el lenguaje tiene como centro el juego de aseverar. De acuerdo con él, las afirmaciones éticas adquieren su significado por la misma vía que las descriptivas, y el contraste entre ambos tipos de afirmaciones responde a la diversidad de aplicaciones funcionales que presenta el juego de aseverar (Price, 2013: 33). Como se ha señalado en el apartado 4, Chrisman ofrece un modo de explicar en virtud de qué el juego de aseverar tiene distintas aplicaciones funcionales. La especificidad de la función normativa que desempeñan las afirmaciones éticas vendría dada por el papel inferencial específico que desempeñan estas afirmaciones. En otras palabras, es posible responder en términos de relaciones inferenciales a la pregunta que inaugura el expresivismo ético: la acción que un hablante realiza por medio de una afirmación puede ser identificada atendiendo a las relaciones inferenciales que la afirmación conlleva. Así, la tesis de que algunas afirmaciones describen estados de cosas mientras que otras expresan intenciones de los hablantes es compatible con una explicación homogénea de la semántica de ambas clases de afirmaciones.



TERCER PREMIO: Inferencialismo en el discurso ético

Tras haber optado por el modelo orgánico de individuación de las proposiciones, conviene revisar una idea que se ha aceptado en las secciones previas. En el apartado 3 se ha defendido que los predicados de orden superior no contribuyen a la proposición expresada por la afirmación en la que aparecen. Esta tesis choca con el modelo inferencialista de la individuación, que lleva a concluir que “Es bueno compartir” expresa una proposición a pesar de que “compartir” no lo hace. Dado que “compartir” no puede actuar como premisa o conclusión de una inferencia y que “Es bueno compartir” sí puede hacerlo, ha de ser que el contenido de “es bueno” contribuye a la proposición expresada por la afirmación. La idea de que los predicados de orden superior no contribuyen a la proposición expresada por la afirmación en la que aparecen se planteó como una consecuencia de la tesis de que las expresiones funcionales de orden superior no se usan para describir el mundo. Tras optar por el inferencialismo, debemos renunciar a inferir que “bueno” no contribuye a la proposición expresada a partir de la tesis de que su función no es describir, puesto que, de acuerdo con la teoría inferencialista, ninguna expresión es significativa en virtud de que representa o contribuye a representar estados de cosas⁴. Si la posición que aquí se ha defendido es correcta, algunas proposiciones son estrictamente normativas. En lo que respecta al predicado de verdad, los partidarios del modelo orgánico pueden conservar la tesis de que este predicado no contribuye al contenido proposicional de las afirmaciones en las que aparece, pero han de hacerlo atendiendo a las relaciones inferenciales que estas afirmaciones conllevan, no a los estados de cosas que supuestamente representarían.

7. Conclusiones

El propósito general de este trabajo ha sido lograr una explicación del discurso ético coherente con los argumentos de Moore y Geach, esto es, una explicación que diera cuenta tanto del comportamiento específico de los predicados éticos como de los rasgos que el discurso ético comparte con el discurso descriptivo. El objetivo específico ha sido contribuir a la explicación inferencialista del significado de las afirmaciones éticas.

La primera parte del trabajo ha consistido en una defensa de las explicaciones anti-representacionistas del significado de los predicados éticos. Se ha argumentado que el comportamiento peculiar de estos predicados es análogo al comportamiento de otros predicados que son funciones de orden superior y que este hecho constituye una razón para suponer que también los predicados éticos son funciones de orden superior.

⁴ Se puede afirmar que el significado de un predicado como “amargo” viene dado por la función descriptiva que desempeña, siempre y cuando esa función descriptiva sea especificada en términos inferenciales.

En rigor, la tesis de que los predicados éticos son de orden superior es compatible con una explicación representacionista de su significado, ya que es posible mantener que los predicados cuyos argumentos son predicados o proposiciones atribuyen una propiedad a los predicados y proposiciones que los saturan. Sin embargo, la explicación representacionista exige la postulación de propiedades *sui generis*, que puede ser evitada por medio de una explicación inferencialista del significado de los predicados de orden superior. Así, se ha mantenido que los predicados de orden superior son herramientas lógico-semánticas cuyo significado es una función del papel inferencial que desempeñan. Para explicar el significado de las afirmaciones éticas debemos atender a la función inferencial que desempeñan las atribuciones de bondad y de maldad. La explicación inferencialista del significado de los predicados éticos ha constituido un primer paso hacia el rechazo de la tesis de la bifurcación propia del expresivismo tradicional, ya que es una consecuencia de esa explicación que los estados mentales son irrelevantes para la individuación del significado de las afirmaciones éticas.

En principio, la opción por una explicación anti-representacionista del significado de algunas expresiones no obliga a renunciar a una explicación representacionista del significado de otras. Por ello, la segunda parte del trabajo ha perseguido dilucidar la función que las adscripciones de bondad realizan sin cuestionar la explicación representacionista del discurso descriptivo. Se ha argumentado que las adscripciones de bondad constituyen tomas de posición por parte del hablante ante cursos de acción posibles, incluso en aquellos casos en los que la expresión “es bueno” es saturada con expresiones que tienen forma oracional, y se ha concluido que los predicados éticos toman otros predicados como argumentos. Esta conclusión ha constituido un segundo argumento en contra de la tesis de la bifurcación. Si la primera de las críticas a esta tesis cuestionaba la explicación expresivista tradicional de las afirmaciones éticas, esta segunda crítica cuestiona la explicación representacionista del resto de expresiones. La tesis de que los predicados éticos son funciones de predicados hace que sea necesario recurrir a un modelo inferencialista de individuación de las proposiciones, que no es compatible con el representacionismo.



El inferencialismo ofrece respuestas tanto para los problemas del representacionismo como para los del emotivismo. A diferencia del representacionismo, el inferencialismo no necesita postular la existencia de propiedades no naturales para explicar el comportamiento peculiar de los predicados éticos. El inferencialismo permite mantener que la función de estos predicados no es describir, a la vez que la tesis de que sus argumentos son acciones explica la vinculación de las afirmaciones éticas con la acción. Frente al emotivismo, el inferencialismo da cuenta de las similitudes que se observan entre las afirmaciones éticas y las descriptivas.

TERCER PREMIO: Inferencialismo en el discurso ético

El rechazo de la tesis de la bifurcación distancia a la solución inferencialista de los expresivismos locales de Blackburn, Ridge y Bar-On y Chrisman, que conservan el supuesto de que las proposiciones genuinas representan estados de cosas. La idea de que hay que distinguir entre una noción estricta y una noción deflacionaria de proposición para dar cuenta de aquellas expresiones que se comportan como si tuvieran contenido proposicional a pesar de que no representan estados de cosas no es sino la consecuencia de un supuesto no justificado, a saber, que el representacionalismo es una teoría correcta. Los problemas que el discurso ético ha planteado, desde Moore hasta estos expresivismos recientes, surgen como consecuencia del supuesto de que la función fundamental del lenguaje es describir. Si la posición que aquí se ha defendido es acertada, los humanos somos, antes que animales que describen, animales que dan y piden razones.



www.solofici.org

BIBLIOGRAFÍA

- AYER, A.J. (1936/1971) *Language, Truth and Logic*. Londres: Penguin Books.
- BAR-ON, D. y CHRISMAN, M. (2009) "Ethical Neo-Expressivism". En: Shafer- Landau, R. (ed.) (2009) *Oxford Studies in Metaethics*, 4. Oxford, Nueva York: Oxford University Press, 133-165.
- BLACKBURN, S. (1993) *Essays in Quasi-Realism*. Cary, NC, USA: Oxford University Press.
- BRANDOM, R. (2000/2002) *La articulación de las razones: una introducción al inferencialismo*. Madrid: Siglo XXI. [Traducción de Eduardo de Bustos y Eulalia Pérez Sedeño].
- CHRISMAN, M. (2008) "Expressivism, Inferentialism, and Saving the Debate". *Philosophy and Phenomenological Research*, 77, 2, 334-358.
- (2010) "Constructivism, Expressivism and Ethical Knowledge". *International Journal of Philosophical Studies*, 18, 3, 331-353.
- (2011) "Ethical Expressivism". En: Miller, C. (ed.) (2011) *The Continuum Companion to Ethics*. Londres: Continuum, 29-54.
- FRÁPOLLI, M.J. (2013) *The Nature of Truth. An Updated Approach to the Meaning of Truth-Ascriptions*. Dordrecht: Springer.
- (2016) "Biting Geach's Bullet". Manuscrito.
- FRÁPOLLI, M. J. y VILLANUEVA, N. (2012) "Minimal Expressivism". *Dialectica*, 66, 4, 471-487.
- (2013) "Frege, Sellars, Brandom: expresivismo e inferencialismo semánticos". En: Pérez, D. (coord.) (2013) *Perspectivas en la filosofía del lenguaje*. Zaragoza: Prensas de la Universidad de Zaragoza, 583-617.
- (2015) "Expressivism, Relativism, and the Analytic Equivalence Test". *Frontiers in Psychology* [en línea], 6, <http://dx.doi.org/10.3389/fpsyg.2015.01788>.
- FREGE, G. (1879/1972) *Conceptografía*. México: Universidad Nacional Autónoma de México. [Traducción de Hugo Padilla].
- (1891/1984) "Función y concepto". En: (1984) *Estudios sobre semántica*. Barcelona: Ediciones Orbis, 17-48. [Traducción de Ulises Moulines].
- (1892/2005) "Sobre sentido y referencia". En: Valdés, L.M. (ed.) (2005) *La búsqueda del significado*. Madrid: Tecnos, 29-49. [Traducción de Luis Manuel Valdés Villanueva].
- (1918/1998) "El pensamiento: una investigación lógica". En: Valdés, L.M. (ed.) (1998) *Ensayos de semántica y filosofía de la lógica*. Madrid: Tecnos, 196-225. [Traducción de Luis Manuel Valdés Villanueva].
- GEACH, P. (1960) "Ascriptivism". *The Philosophical Review*, 69, 2, 221-225. GIBBARD, A. (2003/2009) *Thinking How to Live*. Cambridge, US: Harvard University Press.
- (2012) *Meaning and Normativity*. Oxford: Oxford University Press.
- MILLER, A. (2003) *An Introduction to Contemporary Metaethics*. Cambridge: Polity. MOORE, G. E. (1903/2002) *Principia Ethica*. Barcelona: Crítica. [Traducción de MaríaVázquez Guisán].
- PRICE, H. (2013) *Expressivism, Pragmatism and Representationalism*. Cambridge: Cambridge University Press.
- RIDGE, M. (2006) "Ecumenical Expressivism: Finessing Frege". *Ethics*, 116, 2, 302-336.
- SCHROEDER, M. (2008) *Being for*. Nueva York: Oxford University Press.
- STEVENSON, C. L. (1937) "The Emotive Meaning of Ethical Terms". *Mind*, 46, 181, 14-31.



SOCIEDAD DE
LÓGICA,
METODOLOGÍA Y
FILOSOFÍA DE LA
CIENCIA EN ESPAÑA

Noviembre de 2017

SOCIEDAD DE
LÓGICA,
METODOLOGÍA Y
FILOSOFÍA DE LA
CIENCIA EN
ESPAÑA

www.solofici.org

Para envíos al boletín:
ablancos@ucm.es

www.solofici.org

