

REVISTA DISCUSIONES FILOSÓFICAS

ISSN 0124-6127

- Fundada en enero de 2000 -

Periodicidad semestral

Tiraje 300 ejemplares

Año 11, No 17, 294 p.

julio - diciembre, 2010

Manizales - Colombia

Rector

Universidad de Caldas

Ricardo Gómez Giraldo

Vicerrector Académico

Germán Gómez Londoño

Vicerrector de Investigaciones y Postgrados

Carlos Emilio García Duque

Vicerrector Administrativo

Fabio Hernando Arias Orozco

Vicerrectora de Proyección

Fanny Osorio Giraldo

DISCUSIONES FILOSÓFICAS

Es una publicación del Departamento de Filosofía de la Universidad de Caldas. Sus temas de interés son la Filosofía (en todas sus clasificaciones), y la Literatura (mirada desde un punto de vista filosófico). Nuestro objetivo es difundir resultados de investigación, propiciar el debate crítico sobre las tesis planteadas en los artículos, y contribuir a la literatura y la crítica filosófica mediante recensiones y traducciones sobre los temas ya enunciados. La revista está dirigida a estudiantes, investigadores, profesores y profesionales en filosofía y literatura, así como a otros lectores que tengan afinidad por nuestros temas de interés.

Los artículos que aparecen en esta revista están indexados en el *Philosopher's Index*
página web:

<http://www.ucaldas.edu.co/vinvestyp/index.php?option=content&task=view&id=46>

La revista Discusiones Filosóficas está indexada en Pubindex categoría B

Scientific Electronic Library On-line SciELO

Ulrich's Periodicals Directory

COMITÉ EDITORIAL

Director

Carlos Emilio García Duque

Julián Serna Arango (Universidad Tecnológica de Pereira)

Francisco Cortés Rodas (Universidad de Antioquia)

Juan Manuel Jaramillo (Universidad del Valle)

Alejandro Patiño Arango (Universidad de Caldas)

COMITÉ CIENTÍFICO INTERNACIONAL

Carlos Ulises Moulines (Institut für Philosophie Logik und Wissenschaftstheorie.

Universidad de München - Alemania)

Garrett Thomson (College of Wooster - USA)

Andrés Rivadulla Rodríguez (Universidad Complutense de Madrid - España)

Anandi Hattiangadi (Saint Hilda's College Oxford - Inglaterra)

Décio Krause (Universidad Federal de Santa Catarina - Brasil)

Luis Salvático (Universidad Nacional de Córdoba - Argentina)

COMITÉ TÉCNICO DE APOYO A LA EDICIÓN

Juan David Giraldo Márquez - Coordinación

Juan David López González - Diagramación

Carlos Eduardo Tavera Pinzón - Soporte Técnico

Carlos Fernando Nieto Betancur - Implementación Metodología SciELO

Carlos Emilio García Duque - Traducción de los resúmenes al inglés

Laura Rueda Chaparro - Corrector de estilo

Pedro Antonio Rojas Valencia - Monitor

VENTAS, SUSCRIPCIONES Y CANJES

Vicerrectoría de Investigaciones y Postgrados

Universidad de Caldas - Sede Central

Calle 65 No 26 - 10

Apartado Aéreo: 275

Teléfonos: (+6) 8781500 ext. 11222

E-mails: vinves@ucaldas.edu.co

revistascientificas@ucaldas.edu.co

discufilo@ucaldas.edu.co

Manizales - Colombia

EDICIÓN

Universidad de Caldas

Vicerrectoría de Investigaciones y Postgrados



Imagen Portada:
Bas C. Van Fraassen

La responsabilidad de lo expresado en cada artículo es exclusiva del autor y no expresa ni compromete la posición de la revista.

ph

CONTENIDO CONTENTS

PRESENTACIÓN EDITORIAL <i>EDITORIAL</i>	7
"RELATIONAL QUANTUM MECHANICS". ROVELLI'S WORLD <i>"MECÁNICA CUÁNTICA RELACIONAL". EL MUNDO DE ROVELLI</i> <i>Bas C. van Fraassen</i>	13
A LITTLE GIVE AND TAKE: PROBLEMS IN THE EMPIRICISM OF SELLARS AND HIS FOLLOWERS <i>UN POCO DE TOMA Y DACA: PROBLEMAS EN EL EMPIRISMO</i> <i>DE SELLARS Y SUS SEGUIDORES</i> <i>Michael Davis</i>	53
CONCEIVING THE IMPOSSIBLE AND THE MIND-BODY PROBLEM <i>CONCEBIR LO IMPOSIBLE Y EL PROBLEMA MENTE-CUERPO</i> <i>Thomas Nagel</i>	69
MODUS VIVENDI, OVERLAPPING CONSENSUS AND STABILITY <i>MODUS VIVENDI, CONSENSO COINCIDENTE Y ESTABILIDAD</i> <i>John Rudisill</i>	87
FREGE AND NUMBERS AS SELF-SUBSISTENT OBJECTS <i>FREGE Y LOS NÚMEROS COMO OBJETOS AUTO-SUBSISTENTES</i> <i>Gregory Lavers</i>	97
THE LIAR AND THE NEW T-SCHEMA <i>EL MENTIROSO Y EL NUEVO ESQUEMA T</i> <i>Stephen Read</i>	119
LOS DEÍCTICOS. UN PROBLEMA PARA LA SEMÁNTICA DE GOTTLOB FREGE <i>INDEXICALS. A PROBLEM FOR GOTTLOB FREGE'S SEMANTIC</i> <i>Ángela Rocío Bejarano Chaves</i>	139

JOHN RAWLS: A LIBERDADE DOS ANTIGOS E A LIBERDADE DOS MODERNOS – PRESSUPOSTOS DA JUSTIFICAÇÃO COERENTISTA. <i>JOHN RAWLS: THE LIBERTY OF ANCIENTS AND THE LIBERTY</i> <i>OF MODERNS-PRESUPPOSITIONS OF COHERENTIST JUSTIFICATION</i> <i>Elnora Gondim</i>	151
CRÍTICAS AL ENFOQUE CAUSAL DE LA ACCIÓN PROPUESTO POR HARRY FRANKFURT <i>CRITICISMS TO THE CAUSAL APPROACH OF THE ACTION PROPOSED BY HARRY FRANKFURT</i> <i>Jorge Gregorio Posada Ramírez</i>	167
PRAGMÁTICA SIN PRAGMATISMO <i>PRAGMATICS WITHOUT PRAGMATISM</i> <i>Gabriel Méndez Hincapié</i>	181
LA METAFILOSOFÍA DE LOS PRAGMATISTAS CLÁSICOS <i>THE METAPHILOSOPHY OF CLASSIC PRAGMATISTS</i> <i>Gregory Fernando Pappas</i>	205
LA NEUTRALIDAD VALORATIVA DE LA CIENCIA Y EL CONFLICTO ENTRE LA RAZÓN Y LA PASIONES <i>SCIENCE'S NEUTRALITY AND THE CONFLICT BETWEEN REASON AND PASSIONS</i> <i>Carlos Emilio García y Pablo Rolando Arango</i>	223
PACTO SOCIAL Y LIBERTAD EN THOMAS HOBBS <i>SOCIAL PACT AND LIBERTY IN THOMAS HOBBS</i> <i>John Fredy Lenis Castaño</i>	239
LA TEMPOREIDAD EXTÁTICO-HORIZONTAL COMO ORIGEN DE LA TRASCENDENCIA DEL DASEIN <i>ECSTATIC-HORIZONTAL TEMPORALITY AS THE ORIGIN OF DASEIN'S TRASCENDENCY</i> <i>Carlos Di Silvestre</i>	255
RESEÑA CREER EN LA EDUCACIÓN POR VICTORIA CAMPS C. EDITORIAL PENÍNSULA, BARCELONA, 2008. <i>Katherine Esponda Contreras</i>	276
COLABORADORES <i>COLLABORATORS</i>	283

PRESENTACIÓN EDITORIAL

EDITORIAL

Con orgullo y complacencia entregamos a la comunidad filosófica nacional e internacional un nuevo fascículo de nuestra revista *Discusiones Filosóficas*. En este número hemos logrado reunir catorce artículos de investigadores de talla internacional y jóvenes promesas colombianas. Los temas abarcan la filosofía de la ciencia, la filosofía del lenguaje, la filosofía política y moral y la historia de la filosofía. Para comenzar hemos incluido un artículo de Bas van Fraassen sobre mecánica cuántica relacional. En este trabajo van Fraassen examina un texto de Carlo Rovelli y nos explica en que medida su contribución al área logra ofrecernos una nueva visión de cómo es el mundo de la mecánica cuántica. Como se sabe, Rovelli también propone un programa para derivar el formalismo de la teoría de un conjunto de postulados simples que pertenecen al procesamiento de la información. Van Fraassen advierte que se concentrará totalmente en la exploración del mundo de la mecánica cuántica tal como lo representa Rovelli, y que mostrará en qué medida ese mundo resulta fascinante, en parte debido a la dependencia de Rovelli sobre el enfoque de la teoría de la información para los fundamentos de la mecánica cuántica, y en parte debido a que su presentación implica asumir una postura en una parte fundamental de la filosofía misma. Nuestro segundo artículo “Un poco de toma y daca: problemas en el empirismo de Sellars y sus seguidores” del conocido filósofo Michael Davis, examina el rechazo del empirismo fundacionalista como se aprecia en la discusión que hace Sellars del *mito de lo dado*. Según Davis, Sellars ataca el mito desde dos ángulos principales que corresponden a los dos elementos del empirismo: la idea de que nuestras creencias son justificadas por el mundo, y la idea de que nuestros conceptos se derivan de la experiencia. Aunque Sellars ataca la segunda, de manera correcta, también ataca la primera, de manera incorrecta. Por ende, Sellars rechaza la idea del sentido común según la cual, por lo menos algunas de nuestras ideas se justifican por apelación al mundo empírico. Davis intenta descubrir las suposiciones que llevan a Sellars a este punto, y determinar como operan las mismas suposiciones en autores como Brandom, Rorty y McDowell, además de ilustrar cómo un rechazo de dichas suposiciones nos dan una manera de superar este problema al no recaer en el fundacionalismo.

A continuación hemos incluido el artículo “Concibiendo lo imposible y el problema mente-cuerpo” de Thomas Nagel. En este trabajo Nagel examina las intuiciones basadas en la perspectiva de la primera persona y muestra cuan fácilmente nos pueden inducir a error sobre lo que es y no es concebible. Este punto usualmente se presenta como apoyo de posiciones reduccionistas comunes sobre el problema mente-cuerpo, pero Nagel lo separa de tal perspectiva. Según el autor, la fuerte apariencia de contingencia en la relación entre el funcionamiento del organismo físico y la mente consciente –una apariencia que depende directa o indirectamente de la perspectiva de la primera persona- tiene que ser una ilusión. Nagel considera que hay una conexión necesaria en

ambas direcciones entre lo físico y lo mental, pero que no puede descubrirse *a priori*. La opinión se divide fuertemente entre la credibilidad de algún tipo de reduccionismo funcionalista, pero Nagel se ubica del lado antireduccionista de este debate. Nagel concluye que nuestra falta de habilidad para obtener una concepción inteligible de la relación mente-cuerpo es un signo de la inadecuación de nuestros conceptos actuales, y que se requiere algún grado de desarrollo. Luego se incluyó un trabajo de John Rudisill sobre filosofía política. En su “*Modus vivendi, overlapping consensus, and stability*” Rudisill muestra en qué medida la teoría política de Hobbes se puede interpretar como comprometida con alguna forma de neutralidad. Al reconocer el rol que juega la neutralidad en el pensamiento de Hobbes se logra ver que un requisito de neutralidad no es exclusivo del liberalismo. Más allá de esto, sin embargo, Rudisill intenta mostrar que considerar a Hobbes en tal contexto revela ciertos puntos útiles de comparación con el trabajo tardío de Rawls, que plantea preocupaciones sobre la viabilidad de su liberalismo político. Rudisill plantea que el liberalismo político de Rawls, mientras no ofrece una solución de *modus vivendi* a la justificación política, no está bien dotado para asegurar la estabilidad.

Nuestro fascículo prosigue con un artículo de Gregory Lavers titulado “*Frege and numbers as self-subsistent objects*”. En este artículo se argumenta que Frege no asume la posición de metafísico platónico sobre matemáticas que se le atribuye normalmente. Lavers muestra que el proyecto fregeano pasa por dos etapas distintas: la identificación de lo que es verdadero en nuestras nociones ordinarias, y luego la provisión de una explicación sistemática que comparte los aspectos identificados, y que ninguna de las dos etapas involucra mucha metafísica. El autor critica en detalle la interpretación que hace Dummett de los párrafos §§55-61 del *Grundlagen*. Estas secciones están bajo el encabezado ‘Todo número es un objeto auto-subsistente’ y Dummett las describe como las que contienen los peores argumentos planteados por Frege. Lavers plantea que, esencialmente, todos los puntos interpretativos de Dummett son erróneos. El autor finaliza su análisis demostrando que los planteamientos de Frege sobre la independencia de las matemáticas con respecto a los humanos y sus actividades tampoco lo comprometen con ninguna posición metafísica particular. Acto seguido hemos incluido el trabajo “*The liar and the new t-schema*” de Stephen Read. En este artículo se examina la costumbre de considerar que toda instancia del esquema T es verdadera a partir de la publicación del estudio de Tarski sobre el concepto de *verdad* en los años 30. Sin embargo, Read nos recuerda que algunas instancias del esquema, como las ejemplificadas por la oración del mentiroso, son falsas. El autor demuestra que un esquema mejor permite un tratamiento uniforme de la verdad en el que las paradojas semánticas resultan ser simplemente falsas. En esta misma área de trabajo, hemos incluido el artículo de Ángela Bejarano titulado “*Los deícticos. Un problema para la semántica de Gottlob Frege*”. En este texto Bejarano plantea que hay razones suficientes para considerar que los deícticos plantean un problema para la propuesta semántica de Gottlob Frege. La autora aborda su proyecto en dos secciones. En la primera, expone el programa semántico del lógico alemán por medio de tres tesis estructurantes. En la segunda, introduce la cuestión de los deícticos, y explora en qué medida representan un problema para dicho programa.

Tras estos análisis en la filosofía del lenguaje, hemos incluido un texto de Elnora Gondim titulado “*John Rawls: a liberdade dos antigos e a liberdade dos modernos - pressupostos da justificação coerentista*”. En este artículo se plantea como el énfasis en la libertad individual y la igualdad de todos los ciudadanos que se ve en los dos

principios rawlsianos de justicia hace que esa justicia como equidad establezca una unión entre la libertad individual y la colectiva como puede apreciarse explícitamente en las dos facultades morales contenidas en la concepción de persona de la teoría de Rawls, es decir, en la idea de la racionalidad y en la concepción de razonabilidad. De acuerdo con Gondim, Rawls elabora así una unión entre lo público y lo privado a través de la cual inserta la antigua tradición de libertad en la justicia como equidad. En la justicia como equidad, el uso del método del equilibrio reflectivo, la autonomía, la libertad, el concepto normativo de persona, el procedimentalismo puro, el orden lexicográfico atribuido al primer principio, a la concepción política de justicia y a la base pública de justificación, muestra que Rawls considera la libertad de los antiguos comparable a la de los modernos y, sobre la base de este argumento, evidencia que hay un fuerte apoyo conceptual de corte coherentista con respecto a su modelo de justificación.

A continuación se incluyó el trabajo “Críticas al enfoque causal de la acción propuesto por Harry Frankfurt” de Jorge Gregorio Posada. El profesor Posada intenta mostrar que la crítica de Frankfurt a lo que él denomina enfoque historicista causal de la acción es injustificada. Según el autor, una teoría de la acción que distinga a las acciones de los sucesos a partir de su historia causal, no implica que la explicación de la acción se desplace a algo que ocurrió antes de la acción, su causa, o que descuide en la explicación de lo que es una acción, la acción misma. La tesis resulta más clara al comprender que Davidson y Searle, quienes son asumidos por Frankfurt como historicistas causales, si bien parten de la noción de causa como esencial para explicar lo que es una acción, no suponen que debe repararse en algo que sucede antes de la acción para explicarla, o que la causa de una acción sea un suceso que necesariamente la antecede.

El presente fascículo prosigue con dos trabajos sobre pragmatismo. El primero de ellos, “Pragmática sin pragmatismo” de Gabriel Méndez Hincapié, muestra la manera en que se relacionan algunos aspectos pragmáticos en la obra de Wittgenstein con elementos centrales del pragmatismo de Peirce. En su análisis, Méndez opta por una vía alternativa que supera el dualismo wittgensteniano al partir de un esquema semiótico-modal genuinamente triádico, es decir, no-dualista. Méndez pretende mostrar que dos herramientas auténticamente triádicas, a saber, los *gráficos existenciales* y la *semiótica universal* de Peirce, integradas en un modelo de “lógica topológica”, constituyen un instrumental válido y original para re-dimensionar la noción de ‘juego de lenguaje’. En el segundo trabajo, “La metafilosofía de los pragmatistas clásicos” de Gregory Fernando Pappas, se arguye que la metafilosofía de los pragmatistas constituye su contribución más importante a la historia de la filosofía y su característica distintiva. Como se sabe, los filósofos clásicos americanos (Peirce, James, y Dewey) y los filósofos pragmatistas hispanos, Ortega y Gasset y Risieri Frondizi, le otorgaron un punto de partida práctico a la filosofía. Tras explicar qué significa sostener que la experiencia es el punto de partida de la filosofía, Pappas examina las razones que tienen estos filósofos para sostener que la experiencia es el punto de partida apropiado si se quiere que la filosofía sea empírica y relevante.

Enseguida se incluyó el artículo “La neutralidad valorativa de la ciencia y el conflicto entre la razón y las pasiones” de Carlos E. García y Pablo Arango. En este trabajo se discute, en primer lugar, la tesis de la neutralidad moral de la ciencia. Los autores examinan, a la luz de los puntos de vista del Positivismo Lógico, los posibles significados de esta tesis, así como la plausibilidad de cada uno. Concluida esta tarea, ellos analizan el problema de las relaciones entre la ciencia y el mal. Apelando a algunas tesis de

David Hume, argumentan que el problema no reside en la ciencia como tal, sino más bien en la parte irracional de la naturaleza humana. García y Arango concluyen que lo que se necesita para enfrentar las enormes amenazas destructivas que surgen de ciertas aplicaciones de la ciencia, no es un cambio en el aspecto racional del hombre, sino justamente un cambio en su parte irracional. Por otra parte, el profesor John Fredy Lenis aportó el artículo “Pacto social y libertad en Thomas Hobbes”. En dicho trabajo Lenis procura dilucidar algunas implicaciones antropológico-pedagógicas del pacto instituidor del Estado civil propio de la teoría política de Thomas Hobbes. Según Lenis tales implicaciones están relacionadas, por un lado, con el sentido educativo de la instancia jurídico-penal y, por otro, con el tipo de libertad a la que pueden aspirar los ciudadanos en dicho Estado. Para ilustrar esta tesis, el autor examina: el pacto como instituidor del Estado civil; el sentido educativo del Estado; el concepto de libertad civil; y, finalmente, la ética del soberano y los límites del Estado.

Los artículos de este fascículo se cierran con el trabajo “La temporeidad extático-horizontal como origen de la trascendencia del *Dasein*” de Carlos Di Silvestre. En este trabajo el autor examina un aspecto relevante de la interpretación heideggeriana del tiempo como temporeidad del *Dasein*: su determinación como horizonte trascendental de la comprensión del ser en general. Según Di Silvestre, si bien es cierto que Heidegger abandona el proyecto ontológico de una interpretación *temporaria* del sentido del ser (tal como aparece en *Ser y Tiempo*), es posible mostrar que, en textos posteriores, mantiene y profundiza su concepción de la temporeidad como horizonte trascendental de comprensión. Esta tesis se refuerza mediante una reconstrucción de la noción de *esquema horizontal* de los éxtasis tempóreos, a fin de aclarar en qué sentido el tiempo constituye el origen ontológico del ser-en-el-mundo (trascendencia) como ámbito abierto de manifestación significativa.

A partir de esta edición se registran algunos cambios en nuestro comité científico internacional y en nuestro comité editorial. Con enorme regocijo le extendemos un cordial saludo de bienvenida a los doctores Andrés Rivadulla Rodríguez, de la Universidad Complutense de Madrid, al doctor Anandi Hattiangadi del St Hilda's College Oxford, al doctor Décio Krause de la Universidad Federal de Santa Catarina y al doctor Francisco Cortés Rodas de la Universidad de Antioquia, quienes amablemente han aceptado participar en nuestro proyecto editorial como miembros de nuestros distintos comités. También queremos agradecer el apoyo incondicional de los doctores Fredy Tellez, Valentina Marulanda y Liliana Herrera quienes nos acompañaron por algo más de diez años en esta aventura. Gracias a su invaluable apoyo la revista ha alcanzado el desarrollo que hoy tiene y se perfila como una de las mejores revistas de filosofía en el país. No quisiera dejar pasar la oportunidad sin agradecerle también a Raúl Andrés Jaramillo, quien durante su estadía en España, realizó importantes contactos con potenciales autores y árbitros de nuestras dos últimas ediciones. A sus buenos oficios debemos buena parte de las contribuciones que se publican en esta entrega. Una vez más quiero invitar a los colegas colombianos y de otros países a que nos hagan llegar sus comentarios y propuestas de publicación a la dirección electrónica de la revista.

El Editor

ph

DISCUSIONES FILOSÓFICAS

“RELATIONAL QUANTUM MECHANICS”. ROVELLI’S WORLD

“MECÁNICA CUÁNTICA RELACIONAL”. EL MUNDO DE ROVELLI

BAS C. VAN FRAASSEN
San Francisco State University. discufilo@ucaldas.edu.co

RECIBIDO EL 2 DE SEPTIEMBRE DE 2010 Y APROBADO EL 30 DE SEPTIEMBRE DE 2010

RESUMEN

El inspirador *Relational Quantum Mechanics* de Carlo Rovelli cumple varios propósitos de manera simultánea: proporciona una nueva visión de cómo es el mundo de la mecánica cuántica y ofrece un programa para derivar el formalismo de la teoría de un conjunto de postulados simples que pertenecen al procesamiento de la información. En este artículo propongo que nos concentremos totalmente en lo primero, para explorar el mundo de la mecánica cuántica tal como lo representa Rovelli. Es un mundo fascinante, en parte debido a la dependencia de Rovelli sobre el enfoque de la teoría de la información para los fundamentos de la mecánica cuántica, y en parte debido a que su presentación implica asumir una postura en un parte fundamental en la filosofía misma.

PALABRAS CLAVE

Carlo Rovelli, Einstein-Podolski-Rosen, información cuántica, mecánica cuántica relacional.

ABSTRACT

Carlo Rovelli’s inspiring *Relational Quantum Mechanics* serves several aims at once: it provides a new vision of what the world of quantum mechanics is like, and it offers a program to derive the theory’s formalism from a set of simple postulates pertaining to information processing. I propose here to concentrate entirely on the former, to explore the world of quantum mechanics as Rovelli depicts it. It is a fascinating world in part because of Rovelli’s reliance on the information theory approach to the foundations of quantum mechanics, and in part because its presentation involves taking sides on a fundamental divides within philosophy itself.

KEY WORDS

Carlo Rovelli, Einstein-Podolski-Rosen, quantum information, relational quantum mechanics

Rovelli's inspiring "Relational Quantum Mechanics" provides an original vision of what the world of quantum mechanics is like¹. It is fascinating in part because its presentation involves taking sides on a fundamental divide within philosophy itself.

1. PLACING ROVELLI

1.1 Rovelli's description of roveli's world

In Rovelli's world there are no observer-independent states, nor observer-independent values of physical quantities. A system has one state relative to a given observer, and a different state relative to another observer. An observable has one value relative to one observer, and a different value relative to another observer. (The relativity of values of observables follows from the relativity of states in this view, because Rovelli emphatically retains the 'eigenstate-eigenvalue link': observable A has value x precisely if the system to which A pertains is in an eigenstate of A. However, this must be read so as to accommodate 'vague' ascriptions of values, since the states of a system relative to various observers are generally mixed rather than pure). 'Observer' does not have connotations of humanity or consciousness here –each system provides its own frame of reference relative to which states and values can be assigned. The analogy being drawn on continues a convention adopted at the birth of the theory of relativity, where observers were equated to moving spatial frames of reference.

We want to ask at once: what are the absolutes, the invariants, the features that do characterize these systems, in ways that are not relative to something else?

That remains crucial to the understanding of this view of the quantum world. Following Rovelli's own convenient fiction of observers who measure and assign states to the objects they measure, we can think of those observers as having assimilated Rovelli's view, and thus having available some of his observer-independent description of what is going on. In assigning a state to a measured object, which includes information about probabilities of outcomes of possible future measurements, the

¹ ROVELLI, C. "Relational Quantum Mechanics". In: *International Journal of Theoretical Physics*, 1996. vol. 35, p. 1637-1678. I will refer by section numbers, since a revised version is available on the web.

observer draws on stable observer-independent features (notably, the algebra of observables and the 'transition probabilities' provided by quantum mechanics).

1.2 History of quantum theory interpretation

We can relate Rovelli's approach to a fundamental division among interpretations of quantum mechanics that was outlined by John Wheeler. When Everett published his seminal paper in 1957, Wheeler added a commentary acknowledging that throughout the history of Quantum Mechanics so far, there had been two views in tension with each other, and he argued that Everett had finally made the 'one true story of the universe' version feasible:

(1) The conceptual scheme of "relative state" quantum mechanics is completely different from the conceptual scheme of the conventional "external observation" form of quantum mechanics and (2) The conclusions from the new treatment correspond completely in familiar cases to the conclusions from the usual analysis. The rest of this note seeks to stress this *correspondence in conclusions* but also this *complete difference in concept*².

Wheeler is here contrasting Everett's conception with the older 'external observation' conception, that he describes as follows:

The "external observation" formulation of quantum mechanics has the great merit that it is dualistic. It associates a state function with the system under study –as for example a particle– but not with the *ultimate* observing equipment. The system under study can be enlarged to include the original object as a subsystem and also a piece of observing equipment –such as a Geiger counter– as another subsystem. At the same time the number of variables in the state function has to be enlarged accordingly. However, the *ultimate* observing equipment still lies outside the system that is treated by a wave equation³.

Rovelli clearly places himself in the older 'external observation' formulation, opposite to the new one that Wheeler lauds. But there

² WHEELER, J. "Assessment of Everett's 'Relative State' formulation of quantum theory". In: Review Mod. Phys., 1957. vol. 29, p. 463.

³ *Ibidem*.

is one very important difference that places Rovelli somewhat nearer Everett's. Rovelli takes seriously the idea that any and every system can play the role of 'ultimate observing equipment':

By using the word "observer" I do not make any reference to conscious, animate, or computing, or in any other manner special, system. I use the word "observer" in the sense in which it is conventionally used in Galilean relativity when we say that an object has a velocity "with respect to a certain observer". The observer can be any physical object having a definite state of motion. For instance, I say that my hand moves at a velocity v with respect to the lamp on my table. Velocity is a relational notion (in Galilean as well as in special relativistic physics), and thus it is always (explicitly or implicitly) referred to something; it is traditional to denote this something as the observer, but it is important in the following discussion to keep in mind that the observer can be a table lamp⁴.

Thus Rovelli insists that all systems "are assumed to be equivalent, there is no observer-observed distinction". In saying this he does not take back his rejection of the notion of observer-independent states or observer-independent values of physical quantities. Instead, he means that just as in his guiding example of relativity theory, every physical object can be taken as defining a frame of reference to which all values of physical quantities are referred. Related to this objectification of the 'external observer' is his conception of *information* in physics:

Also, I use information theory in its information-theory meaning (Shannon): information is a measure of the number of states in which a system can be –or in which several systems whose states are physically constrained (correlated) can be. Thus, a pen on my table has information because it points in this or that direction. We do not need a human being, a cat, or a computer, to make use of this notion of information⁵.

Rovelli takes it that any system can in principle have information about any other, due to a previous interaction, for he equates the having of information in its physical sense with a correlation that has been effected by such an interaction:

⁴ *Ibid.*, end sect. I.

⁵ *Ibidem.*

(...) any physical system may contain information about another physical system. For instance if we have two spin $-1/2$ particles that have the same value of the spin in the same direction, we say that one has information about the other one. Thus observer system in this paper is any possible physical system (with more than one state). If there is any hope of understanding how a system may behave as observer without renouncing the postulate that all systems are equivalent, then the same kind of processes –"collapse"– that happens between an electron and a CERN machine, may also happen between an electron and another electron. Observers are not "physically special systems" in any sense.

We must treat this with some delicacy, since the usual explanation of such correlations or entanglements is in terms of states conceived of as observer-independent. The standard quantum mechanical formalism is used here, but understood in a new way.

Given the comparative loss of popularity of the older 'external observation' approach, at least among those who work on foundations of physics, Rovelli's return to it at this date imparts his view with a stimulating sense of novelty.

1.3 Information-theory approach from groenewold to the present

Noting the emphasis Rovelli puts on information, it is also important to place Rovelli's approach with respect to the information-theory approach. This is a very lively new development. While there were beginnings and precedents, this has recently taken a quite radical turn, and Rovelli's work can be seen as involved in that turn. Let's look at the beginnings first and then at the radical agenda in such recent work as that of Christopher Fuchs, Jeffrey Bub and their collaborators.

In the 1950s H. J. Groenewold advocated that we should regard quantum states as just summaries of information obtained through measurement. There are some striking similarities between Groenewold's description of the quantum mechanical situation and Rovellis.

Groenewold (1952, 1957) proposed a formulation of the theory that would contain all its empirical content without referring to states in any essential way. He derided the idea that quantum states are to be thought of on the model of states in classical mechanics. His formulation

re-appears quite clearly in Rovelli's article, though there in a more general form. The idea is that a situation of interest is to be depicted as the effect of a series of measurements, represented by a series of observables (the ones being measured) interspersed with evolution operators (governing evolution between measurements). The sole real problem to be addressed, according to Groenewold, is this:

given the outcomes of preceding measurements, what are the probabilities for outcomes of later measurements in the series?

The answer is formulated in terms of transition probabilities⁶. In the exposition of Rovelli's specific version below I shall explain and illustrate how that goes.

Groenewold offers an argument to the effect that states are to be regarded as 'subjective' or 'observer-relative', determined by information available. Imagine that each measurement apparatus in the series records its outcome⁷. After the entire series has been concluded, a physicist O inspects those recorded results in some order, and assigns states to the system measured for the times of those outcomes using von Neumann's Projection Postulate recipe (which everyone agrees is fine for such narrowly focused predictive tasks). To begin, O assumes some initial state. Groenewold suggests that in absence of other information that could be the entirely uninformative mixture represented by the identity operator on the space. For time t between times t_1 and t_2 where the state $\rho(t_1)$ is assumed or known the calculation looks like this:

$$\rho(t) = U(t, t_2)K(t_2)\rho(t_1) K(t_2)U(t_2, t)$$

(with a correspondingly longer such series for a longer series of measurements between the initial and final time) where the K s are transition operators, and the time-indexed ρ is the ascribed state; the U s are the normal evolutions while no measurement or other interference occurs.

⁶ Groenewold was not the only one; see for example TEMPLE, G. *The general principles of Quantum Theory*. London: Methuen, 1948.

⁷ See DICKE, R. H. "Quantum measurements, sequential and latent". In: *Foundations of Physics*, 1989. vol. 19, p. 385-395, for an argument about how this is physically possible without disturbance; see further the discussion in my *Quantum Mechanics: An Empiricist View* (Oxford: Oxford University Press, 1991. p. 257-258).

But now what would happen if O (or one of his colleagues) decides on a different order for inspection of the recorded outcomes? For the same times, although having started with the same initial knowledge or assumptions about the system, the assignment of states will be quite different.

There is nothing contentious in this imagined scenario itself. The contentious part is Groenewold’s insistence that no other significance is to be accorded to the assignment of states. They are nothing more than compendia of information assumed, known, or gathered through measurements, and thus determined entirely by a specific history, the ‘observer’s’ history. The truly empirically testable part of the theory, he insists, is contained in the transition probabilities. When they are tested, the convenient calculation starts with an assignment of an initial state, but coherence requires only that some such initial assignment leads to the right predictions –the transition probabilities are independent of the states, they are formulable in terms of the observables⁸.

This insistence, that the states be thought of as playing no other role, is at the heart of the recent innovations in the information theoretic approach. Christopher Fuchs presents the program in its most radical form in his much discussed “Quantum Mechanics as Quantum Information (and only a little more)”:

This, I see as the line of attack we should pursue with relentless consistency: The quantum system represents something real and independent of us; the quantum state represents a collection of subjective degrees of belief about something to do with that system (even if only in connection with our experimental kicks to it). The structure called quantum mechanics is about the interplay of these two things –the subjective and the objective (quant-ph/0205039, p. 5).

He submits that “the quantum state is solely an expression of subjective information –the information one has about a quantum system. It has

⁸ They are often presented as probabilities for transitions between states, because the Projection Postulate is generally taken for granted. In my *Quantum Mechanics: An Empiricist View*, I explained them in an intermediate way: the probability is that of the outcome 1 of a measurement of the observable represented by projection on the vector representing second state, given that the system measured is in the first state. But this can easily be replaced by a formulation in terms of the two observables, which are the projections on the two states. For a nice introductory treatment of the theory entirely in this form we can look to TEMPLE, Op. cit.

no objective reality in and of itself". When asked "information about what?" he replies "the answer is 'the potential consequences of our experimental interventions into nature'" (*Ibid.*, p. 7). But Fuchs also has a precise proposal about how to describe the information-updating process in response to measurement (see further Fuchs, 1998.). Drawing on results, both his own and others, he depicts it as a special case of Bayesian updating of opinion by conditionalization. We have to think here, as in Groenewold's scenario, of an epistemic agent with a pertinent prior state of opinion – a physicist who accepts at least the bare minimum of the quantum theory – reacting to recorded measurement outcomes. There is also, without explicit attention paid, for both Groenewold and Fuchs, a presumed coordination, so that tangible physical operations can be univocally represented in terms of an algebra of observables of a certain sort.

This reliance on a fundamental representation of the physical situation – the coordination – becomes clearest in the important paper by Robert Clifton, Jeffrey Bub, and Hans Halvorson. The physical system is characterized by means of an algebra of observables, taken to be a C^* algebra⁹. But states are just generalized probability functions – more accurately, expectation value functions – defined on this algebra of observables. So far that is similar to the approach in more "realistically" understood foundational treatments. The difference comes in what is added now so as to single out quantum theories. What is added is constraints on information transfer, with the states thought of as information depositories. From the premise that those constraints are satisfied, the basic principles of quantum theory are deduced. As reflection on this result, Bub then argued in his "Why the Quantum?" that "A quantum theory is best understood as a theory about the possibilities and impossibilities of information transfer, as opposed to a theory about the mechanics of non-classical waves or particles"¹⁰.

"Information" is here understood as Groenewold specified, in the technical sense of information theory, as measured classically by the Shannon entropy or by the von Neumann entropy for quantum states. And in "Quantum Mechanics is About Quantum Information", Bub argues that:

⁹ This is a very general framework, which allows for the formulation of many sorts of physical theories, both classical and quantum.

¹⁰ BUB, J. "Why the quantum?" In: *Studies in the history and philosophy of modern physics*, 2004. vol. 35, p. 242.

Quantum mechanics represents the discovery that there are new sorts of information sources and communication channels in nature (represented by quantum states), and the theory is about the properties of these information sources and communication channels. You can, if you like, tell a mechanical story about quantum phenomena (...) but such a story, if constrained by the information-theoretic principles, will have no excess empirical content over quantum mechanics. So the mechanical story for quantum phenomena is like an aether story for electromagnetic fields¹¹.

Bub's answer to the question "Information about what?" is just the same as Fuchs's -though in phrasing that shows his special interest in encryption and decoding.

Note once again that some form of coordination is presumed given, without receiving explicit attention: the measurements and their results are assumed univocally representable in terms of the observables that characterize the system. This points to 'absolute' characteristics of the system, which are not aspects of information gathered about it, but pertain to the system itself. That the system is characterizable in such a way is presupposed when certain operations are classified as, or taken to be, means of gathering information about it. Thus here, as for Groenewold (and equally for Rovelli, as we shall see) there is a divide as well as a link between 'subjective' and 'objective' features of the experimental situation.

2. IS THERE A VIEW FROM NOWHERE?

At first sight Rovelli's treatment of states is not exactly what either Groenewold, Fuchs, or Bub appears to advocate.⁶ Rovelli does bring states into the discussion, but as states that measured objects can have *relative to* the measuring system. At first sight we seem to detect a tension between what Rovelli does and what he tells us it is possible to do. What he calls his Main Observation, motivating the view, is similar to Groenewold's though:

¹¹ BUB, J. "Quantum Mechanics is about quantum information". In: Foundations of Physics, 2004. vol. 35, p. 558.

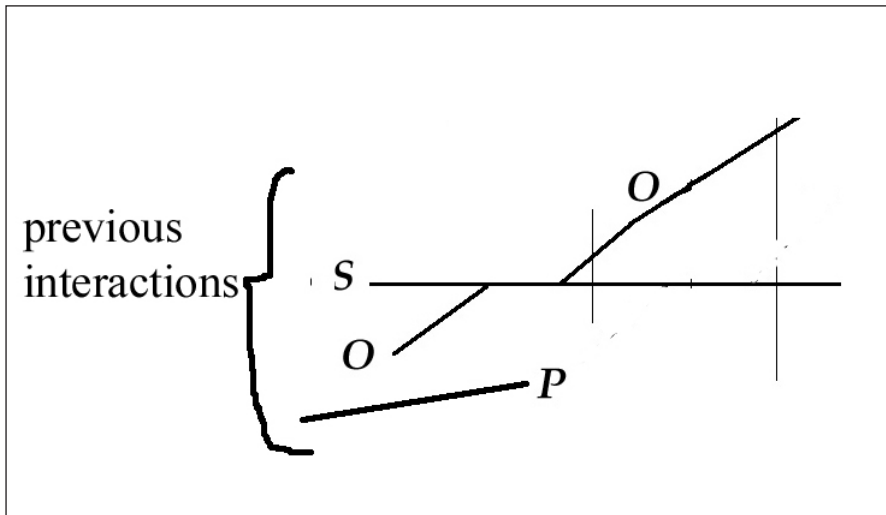
In quantum mechanics different observers may give different accounts of the same sequence of events.

Having rejected the idea of observer-independent states, there is no question of one of those descriptions being the sole truth, with the other illusion or error. Here is an example that Rovelli describes in intuitive terms. I will elaborate on it, in several steps.

Example 1: The two-observer situation

To begin we can characterize it as follows.

There are two observers, O and P, and one other system, S. Observer O measures an observable A on system S, while the second observer P describes this measurement by O on S. (Later on P may make a measurement on S too, or on S+O; but we will leave this unexplored for now).



O registers the value 1, say, and thus assigns pure state $|A, 1\rangle$ to S, or in other words, S is now in state $|A, 1\rangle$ relative to O.

Meanwhile P has the information that this measurement is taking place (presumably on the basis of earlier measurements made on S+O). So P describes O as in an initial state $|init\rangle$ and S+O coupled at the beginning. The state of S + O evolves:

$$(\alpha |A, 1\rangle + \beta |A, 0\rangle) \otimes |\text{init}\rangle \otimes (\alpha |B, 1\rangle + \beta |B, 0\rangle)$$

Here $|B, 1\rangle$ and $|B, 0\rangle$ are the 'pointer reading states' that P uses to characterize observer O when O registers a definite value of 1 or 0 as measurement outcome. That is, the measurement interaction between S and O is such as to effect the requisite correlation between A pertaining to S and B pertaining to O.

If P now wonders what state to assign to S, but does not make a measurement, then he calculates it by the usual 'reduction of the density matrix'. Thus P assigns to S a mixed state, namely the mixture of $|A, 1\rangle$ and $|A, 0\rangle$ in proportions α^2 and β^2 . So we see that O and P assign different states to S. To put it in other words, S has different states relative to O and relative to P.

Rovelli also insists on the orthodox eigenvalue-eigenstate link, so that A takes a value 1 relative to O, but not relative to P –observables have values only relative to observers, and may not have the same value relative to different observers.

But is *this* description of the situation then observer-independent, one that is in fact not relative to any observer? Shouldn't we object that the rationale forbids this, because by Rovelli's lights we can only have descriptions relative to some observer or other?

2.1 General form versus third-observer description

The answer is that there is no incoherence here, but we must carefully distinguish what Rovelli gives us when he presents his view, even in such an example, and the description of the same situation by a third observer. The Example can indeed be elaborated so as to include a third observer, whom we might call ROV. We could imagine that ROV has, on the basis of previous measurements, information that can be summarized by assignments of initial states to O, S, P and their composites relative to ROV, plus later states based on their unitary evolution. We'll look later at how this goes, when we will also have occasion to consider measurements that P can make on O or S later on. But right now we can point out that ROV's information is not to be confused with what Rovelli tells us about this sort of situation. The tension that a reader might feel could be expressed this way:

Rovelli seemingly purports to be giving us a description of the world that would on the one hand be *on the same level as* a description of the rest of the world relative to some given system ROV, and yet on the other hand not relative to anything!

But that is not so at all. Rovelli, who can give these examples, is telling us only something about the *general form* that these observers' descriptions (their information) can take, given that certain measurement interactions have taken place. The resolution of this sensed tension is this: Rovelli does not give any specific such description of the world –he describes *the form that any description which assigns states must take*. Rovelli describes not the world, but the general form of information that one system can have about another– namely as the assignment of states relative to a given system on the basis of information available to that system:

- there is no implication of possible specific information about what there is which is independent of any point of view, but
- there can be knowledge of the form that any such information, relative to a particular vantage point, must take.

So we have here a *transcendental* point of view. Rovelli offers us this knowledge of the general form, the conditions of possibility. We must take very seriously the fact that as he sees it, quantum mechanics is not a theory about physical states, but about ('about'?) information. The principles he sees at the basis of quantum mechanics are principles constraining the general form that such information can take, not to be assimilated to classical evolution-of-physical-state laws.

2.2 The form of an observer's description of the world

This form is constrained by the insistence that specific information, had by one system about another system, can only be a record of actual measurement outcomes. The only way in which there can be information for one observer of what has happened to another observer is through a physical measurement by the former on the latter. Communication, i.e. exchange of information, is physical (cf. end sect III of the article).

Before aiming at greater precision, let's briefly summarize how this happens according to Rovelli's account. A *question* is asked of a system

or source only when an appropriate physical interaction takes place. This interaction is a *measurement* delivering a value for some observable, but also serves as a *preparation*, so that the value obtained has (relative to the theory) predictive content. The probabilities of future measurement outcomes are affected by the outcome obtained –the measured system has gone into a new state relative to the measurement set-up. Thus he accepts (explicitly, in his rejection of the Bohm and modal interpretations) von Neumann's *eigenstate-eigenvalue link*:

the system to which the observable's value pertains is (at that time) in an eigenstate of that observable, corresponding to that value.

But there is a twist, which changes the meaning, so that this says something quite different from its original. The reference is here not to a physical state of the system, but to the state of the system relative to the observer (the measurement apparatus). So the 'collapse' is in that observer's information; the state assigned to the system is a summary of that information.

As mentioned earlier, because of the eigenstate-eigenvalue link it follows that if states are relative, so are values of observables. That an observable takes or has a certain value at a certain moment, that too is observable-relative¹². Because information can only be had by actual, physical measurement, the states assigned will rarely be pure. It is not easy to obtain maximal information about a system, even with respect to targeted observables. So in general the value of an observable, relative to a given observer, will not be sharp.

This information is the subject of two postulates. Let us introduce them in such a way as to spell out what is and is not observer relative. Each physical system S is characterized in the first place by means of a set $W(S) = \{Q_i; i \text{ in } I\}$ of questions that can be asked of it. This association of $W(S)$ with S is not relative to any observer –we may call it the first 'absolute'. Although the presentation differs, this set of questions pertaining to S is essentially the specification of the family of observables that pertain to S . (Eventually, the algebra of observables is reconstructed from this family of questions; for our purposes we need not distinguish

¹² Cfr. LAUDISA, F. & ROVELLI, C. "Relational Quantum Mechanics". In: The Stanford Encyclopedia of Philosophy. [On line] <http://plato.stanford.edu/entries/qmrelational/>. 1995. End of sec. 2.

the two.) When the sets of questions are the same for two systems we may call them *of the same type*.

Secondly, an observer who has been in measurement interaction with a system has a record of the questions that have been asked and the sequence of outcomes thus obtained. That the observer has this *is not relative* to another observer¹³. It is our second 'absolute'. At the same time we must be careful not to equate this fact about the observer with a quantum mechanical state! For while we could try to describe a state that ostensibly is the state that O has if and only if it has a particular sequence of 0s and 1s registered in a series of measurement interactions with S, that would have to be the state of O relative to another observer P who has obtained that information by means of a later measurement on S. We'll see later on whether, or to what extent, there could be a discrepancy, or even a meaningful comparison.

3. STATES AS OBSERVER-INFORMATION

3.1 The postulates constraining information acquisition

POSTULATE 1: (Limited information) There is a maximum amount of relevant information that can be extracted from a system.

Answers to questions have predictive value, but typically, earlier answers become *irrelevant* to the predictions after later answers, *and must do so*. "Irrelevant" and "redundant" are perhaps not entirely apt terms: if a state is to be assigned on the basis of the extracted information, earlier answers must typically have to be *discarded* from the basis on which states are assigned.

For system S there is a definite probability that given question Q will get a yes-answer if asked; this probability can differ for another system of type W(S); moreover, this probability is affected by the answers to previous questions asked.

¹³ By taking this not to be relative, we have in this sequence of 0s and 1s something analogous to Einstein's local coincidences, the 'bed rock' of the representation. Rovelli's criticism of the 'consistent histories' interpretations suggest strongly that he does not allow any ambiguity in this respect.

The 'moreover' establishes that the probabilities in question are transition probabilities. This I will spell out further after the second postulate. Note that what these transition probabilities are is the same regardless of which observer O asks the questions of S . So we have here a third 'absolute'. : But fourthly, if we look at how the probability of future measurement outcomes changes in the course of asking $c = \langle Q_1 Q_2 Q_3 \dots \rangle$ and getting number sequence $sc = [n_1 n_2 n_3 \dots]$ the items that become irrelevant after a certain point *are also the same for all systems of the same type* (the fourth 'absolute'). So, given these notions, we can define:

maximally non-redundant question-answer sequence: one in which no element is irrelevant, but which loses that feature if any question+answer at all is added¹⁴.

Postulate 1 says that this sequence is finite. In a particular case, we can ask for the relevant finite number: how many questions are needed to extract maximal information, leading to the assignment of a pure state relative to the observer? This number *does not depend on which sequence* of questions we pick, and hence also is not relative. Thus Rovelli writes, in a passage immediately following Postulate 1:

One may say that any system S has a maximal "information capacity" N , where N , an amount of information, is expressed in bits. This means that N bits of information exhaust everything we can say about S . Thus, each system is characterized by a number N . In terms of traditional notions, we can view N as the smallest integer such that $N \geq \log_2 k$, where k is the dimension of the Hilbert space of the system S . Recall that the outcomes of the measurement of a complete set of commuting observables, characterizes the state, and in a system described by a $k=2^N$ dimensional Hilbert space such measurements distinguish one outcome out of 2^N alternatives (the number of orthogonal basis vectors): this means that one gains information N on the system.

The number in question therefore depends on the dimension of the state space –if that dimension is finite number k then N is $\log_2 k$ or just above (to make N an integer); the dimension is 2^N , or $(2^N) - 1$. We have a good link here with information theory: the missing information, about

¹⁴ Rovelli introduces and uses the term "complete family sc of information" for "*maximally non-redundant question-answer sequence*".

what this 'source' of type $W(S)$ is like, is extractable in at most N Yes-No questions: the *maximal information capacity* of a system (source) of this type is N bits¹⁵. But now Rovelli adds:

POSTULATE 2 (Unlimited information) It is always possible to acquire new information about a system.

This is not at odds with the first postulate, given that new information can make older information 'irrelevant' (having to be discarded). But it is certainly at odds with the classical ideal of perfectible measurement, as revelation of aspects of the state of the system before measurement, without affecting that state. It entails a certain degree of indeterminism: the maximum possible information at a point does not settle what new information we could get. That is in part because observables can be (totally) *incompatible*: they may have no joint eigen-state:

given a Yes answer to question Q there are many questions Q' such that if they are then asked, their answer *cannot* be Yes with certainty, nor No with certainty.

Notice the modal character of this assertion! In contrast, some questions Q and Q' are *compatible*: on a given occasion, after receiving Yes to Q , the observer has only non-zero probabilities for both possible answers to Q' , but if he then asks Q' , he can base more precise predictions on the fact that he has had these two answers.

New assumption: this indeterminism is not a chaotic randomness, but can be characterized in terms of definite probabilities.

Suppose the first complex apparatus A asks a "complete" question, so it yields a record that provides a *maximally non-redundant question-answer sequence*. Before that question has been asked we have no non-trivial information. Suppose the second apparatus B is equally *complete*, though the question family is very different. Rovelli posits a definite *transition probability* $p(B|A)$ that a Yes answer to B will follow a Yes answer to A , which is both idempotent and symmetric.

¹⁵ Compare: "In particular, I identify one element of quantum mechanics that I would not label a subjective term in the theory; | it is the integer parameter D traditionally ascribed to a quantum system via its Hilbert-space dimension". Chris Fuchs, 2004, Abstract.

Intuitive mnemonics: look at the scenario in which a single source sends many systems of the same type into the series of measurement apparatus for two-valued observables A, B,... that the observer has installed. The stream is diminished by some factor q by the first measurement, then by the transition probability p. Suppose we do A again, then once again the stream is diminished by that factor p. So the number goes from qM to pqM to ppqM by the operations A, AB, ABA so we could write:

$$ABA = pA$$

and this is what a sequence of 1-dimensional projections would do to a vector. It is a way to identify the transition probability. This is numerically equal to the \cos^2 of the angle between the two 1-eigenvectors, onto which they project, or in Hilbert space the squared modulus of the scalar product, or equivalently the trace of the product of the two projections.

After a *maximally non-redundant question-answer sequence* performed by measurement A, the next question might only e.g. ask "is the system in subspace J?", with J of higher dimension –but here there is a definite probability as well, which can be derived (in accordance with the practical calculation suggested by von Neumann's Projection Postulate).

3.2 States as states of information, relative to the observer

Suppose that observer O has put a series of questions to system S and has arrived at the point of attributing $|A, x\rangle$ to S, where x is an eigenvalue of A. Imagine once again a second observer P, whose knowledge (gained earlier through a physical transmission process) was enough to attribute an initial state to S+O, and a Hamiltonian to govern their interaction, enough for him to attribute the evolution in question. Then as we noted above P has the usual 'distant' description of S+O:

- (1) initially it is in state $\sum_i \beta_i (|A, a_i\rangle \otimes |\text{init}\rangle)$
- (2) this evolves into the final state $\sum_i \beta_i (|A, a_i\rangle \otimes |B, a_i\rangle)$

where B is the 'pointer observable' of O -- its value being a recorded sequence of 0s and 1s. Using a reduction, P can attribute a state to S as well, namely

a mixture of states $|A, a_i\rangle$ with weights β_i^2

which is quite different from $|A, 1\rangle$ or $|A, 0\rangle$. According to Rovelli, this is all there is to be said, so far: S has one state relative to O, and another state relative to P. The phrase 'S has state $|A, 1\rangle$ relative to O' means only that the information O has obtained can be summed up or represented by the vector $|A, 1\rangle$. But is the fact that O has certain information a fact that is or is not observer-relative? We must answer this question in the light of two points Rovelli insists on:

- (i) There is no meaning to the state of a system except within the information of a further observer.
- (ii) There is no way a system P may get information about a system O without physically interacting with it, and therefore without breaking down (at the time of the interaction) the unitary evolution description of O.

'Information' has a minimal sense in this context, to say that O has information about S means only that there is a certain correlation in the state of S+O. That much P was able to predict already, and so he can predict something with certainty if a measurement will be made to confirm this. Note that what he is able to predict with certainty amounts to information he already has.

More formally, there is an operator M on the Hilbert space of the S+O system whose physical interpretation is "Is the pointer correctly correlated to A?" If P measures M, then the outcome of this measurement would be yes with certainty, when the state of the S+O system is as in the state described in (2). The operator M is given by

$$\begin{aligned}
 M(|A, 1\rangle \otimes |B, 1\rangle) &= |A, 1\rangle \otimes |B, 1\rangle \\
 M(|A, 1\rangle \otimes |B, 0\rangle) &= 0 \\
 M(|A, 0\rangle \otimes |B, 0\rangle) &= |A, 0\rangle \otimes |B, 0\rangle \\
 M(|A, 0\rangle \otimes |B, 1\rangle) &= 0 \quad (3)
 \end{aligned}$$

where the eigenvalue 1 of M means "yes, the hand of O indicates the correct state of S" and the eigenvalue 0 means "no, the hand of O does not indicate the correct state of S". At time t_2 , the S+O system is in an eigenstate of M with eigenvalue 1; therefore P can predict with certainty that O "knows" the value of A.

Thus, it is meaningful to say, according to the P description of the events E, that O "knows" the quantity A of S, or that O "has measured" the quantity A of S, and the pointer variable embodies the information (cf. middle of section II-D). But of course P had a choice, P could have measured a different observable, say K, to try and find out which result O obtained:

$$\begin{aligned} K(|A, 1\rangle \otimes |B, 1\rangle) &= |A, 1\rangle \otimes |B, 1\rangle \\ K(|A, 1\rangle \otimes |B, 0\rangle) &= 0 \\ K(|A, 0\rangle \otimes |B, 0\rangle) &= 0 \\ K(|A, 0\rangle \otimes |B, 1\rangle) &= |A, 0\rangle \otimes |B, 1\rangle \end{aligned}$$

Intuitively speaking, this is what P would measure to find out what O found. She would get either result 1 or result 0, and would say "O found 1" or "O found 0" accordingly. But can we understand that literally as referring to what O had as information before P made this measurement? If P finds result 1, does that imply that O had found 1 and that O had assigned state $|A, 1\rangle$ to S?

According to Rovelli's rules, *this makes no sense*. An interpretation of quantum measurement as revealing pre-existing values is untenable.

We are now in a position to examine and resolve some puzzles that tend to occur to practically any reader in first acquaintance with this interpretation.

4. PUZZLES POSED AND RESOLVED

All the puzzles will pertain to this basic situation:

O has made a complete measurement on S of two-valued observable A, and has a record of the question asked (call it? A) and the answer received; say 1. Accordingly S has now state $|A, 1\rangle$ relative to O. The pointer observable on O is B, so on the old, pre-Rovelli view one takes it that the existence of the record means that B has value 1. For Rovelli this makes no sense as an observer-independent assertion. To mention values of the pointer observable at all, we need to look at O from the point of view of second observer P.

Meanwhile P had made earlier measurements on O+S and so has the information throughout that this measurement interaction is taking or has taken place. Based on his earlier results and his predictions on that basis, O+S has at the end of the interaction an entangled state, namely $(\beta_0 |B, 0\rangle \otimes |A, 0\rangle) + (\beta_1 |B, 1\rangle \otimes |A, 1\rangle)$, relative to P.

Puzzle 1. Could O and P contradict each other?

Suppose that P will make a measurement on O+S after this point, and later report the result to O. In the meanwhile O makes a prediction with certainty about what P will find. Is it possible that O will find his prediction contradicted by P?

Example: P will measure $(I \otimes A)$ on O+S. P predicts that he will get value 1 with probability < 1 , and value 0 with some probability > 0 . Suppose he gets value 0.

Meanwhile O knows that he has seen value 1, and has a record of that, so assigns himself state $|B, 1\rangle$, and assigns to S the state $|A, 1\rangle$, and therefore to O+S the state $|B, 1\rangle \otimes |A, 1\rangle$. So O predicts with certainty that P's measurement will have result 1. And so O is making a false prediction here, one that is falsified by what P finds.

Reply: The reasoning is questionable in several ways.

To begin we may note an unwarranted assumption in the second paragraph: that O has a state relative to itself here. There was no self-measurement in the story. The relative states are only assigned as summaries of what the real measurement results have been. So as far as this story goes, O has here no state relative to itself, nor does O+S have a state relative to O.

Nevertheless, we can leave aside the issue of whether the possibility of self-measurement could be added to Relational Quantum Mechanics, for there is a much more important point to be made¹⁶.

¹⁶ My suggestion is that this should not be added as a possibility; there certainly seems to me to be no warrant in Rovelli's interpretation for doing so. For a contrary view and a recent 'Wigner's friend' type example presented to challenge information-theoretic approaches (specifically Jeffrey Bub's recent work) see HAGAR, A. & HEMMO, M. "Explaining the unobserved-why QM isn't only about information". In: *Foundation of Physics*, 2006. vol. 36, p. 1295-1324.

The more important point is this. It is not to be assumed that P will ever find 0 in the case in which O has found 1. The insinuation in the above puzzle is that, if this were so, then P’s probabilities would be wrong –and since these probabilities come from quantum mechanics, that such a scenario would contradict quantum mechanics. But this threat disappears as soon as we take heed of what P’s probabilities are. They are what he calculates on a basis that includes no information about what O found. These probabilities would be tested by placing P very often in a situation that matches the information he has. P’s probabilities are for his finding value 1 or value 0 in a situation of *that sort* –where this sort is not identified in terms of what O finds during the process, but only in terms derivable from preceding measurements of O and S that established that an A measurement would take place. We can be sure that if quantum mechanics is right, and P enters into many such situations, he will find values 1 and 0 with the correct frequencies.

What we may note in addition (and to this we will return) is that any immediate attempt to check by measurement whether O’s and P’s outcomes were the same, would get a positive result.

Puzzle 2. But what about ‘immediate repetition’ of measurement?

As von Neumann emphasized, O will predict with certainty that a measurement of A on S, immediately after his own, will find the same value. So does that not apply here, to an immediately subsequent measurement by P?

Reply: No; in Rovelli’s account the collapse of the wave packet appears only in the states relative to a given observer. So his echo of von Neumann is that O will predict with certainty that if *he himself*, or an observer with exactly the same interaction history with a system of type $W(S)$, makes an immediate new measurement of A on S, the same value 1 will appear again.

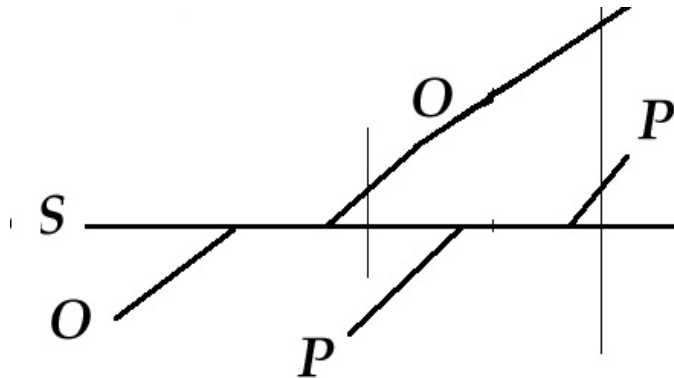
As Rovelli emphasizes, O can get to know P’s result only through a relevant interaction with P, in effect a measurement by O on P. So O could ask the question: “what did P see, when he measured A on S after me?” In the sense that O can measure P’s pointer observable afterward, and get some value. As usual, we cannot assume that the result that O gets is the value that this observable had before O’s measurement.

So does this mean that O and P have no way to find out what either of them saw earlier, as opposed to what it seems now that they have seen? That would still seem almost as puzzling. To answer this properly, we need to construct our puzzle situation with more precision and care.

Puzzle 3. Can an observer find out what was observed earlier on?

To see how we can get into a confusion here, I am going to present this third version of the puzzle first of all in the 'old' style, assuming that states are observer-independent. Then the puzzle will again be resolved by seeing how the understanding of this situation changes on Rovelli's conception.

Let the measured system S start off in a superposition $\sum \beta_i |A, i\rangle$ of eigenstates of an observable A corresponding to distinct eigenvalues, and let us measure A twice, using two measuring systems O and P¹⁷. For simplicity I'll take A to be time-independent (we could put in evolution operators, as Groenewold and Rovelli indicate, but it would not seriously affect the argument), and take the pointer observables of both O and P to be the same observable B.



Then, under the familiar idealized assumptions of a von Neumann measurement, the combined system S+O will be in dynamic state $\sum \beta_i (|A, i\rangle \otimes |B, i\rangle)$ at the end of the first measurement. At the end of the second measurement the dynamic state of S+O+P will be, ignoring phase factors, $\sum \beta_i (|A, i\rangle \otimes |B, i\rangle \otimes |B, i\rangle)$.

¹⁷ Assume that I and O each evolve freely after their measurement interaction, that there is no interaction between O and P, and that both A and the 'pointer-reading observable' B for O commute with the free Hamiltonians for S and O respectively.

By reduction, we have states also for parts of the total system. Write $P[|A,i\rangle]$ for the projection on the ray containing $|A,i\rangle$, etc. At the end of the first measurement, the individual systems S and O are in dynamic states $[\sum |\beta_i|^2 P[|A,i\rangle]]$ and $\sum |\beta_i|^2 P[|B,i\rangle]$, respectively. The final dynamic state of O+P is $\sum |\beta_i|^2 P[|B,i\rangle \otimes |B,i\rangle]$. Following von Neumann, *assuming collapse*, we reason as follows:

As for the individual states, because S and O interact by a measurement interaction, S ends up in some $|A,k\rangle$, with O in the corresponding $|B,k\rangle$. At the conclusion of the first measurement, the pointer reading observable B on O thus has the value k as well, we will say that its pointer reads k. Likewise, at the end of the second measurement, S ends up in some state $|A,m\rangle$ with P in the corresponding $|B,m\rangle$; its pointer the reads m. **Moreover, $m=k$.**

Suppose we want to check now whether that is so. Then we can have a **third** measurement, of that 'agreement observable' that Rovelli describes, as follows:

Let M be an observable for the combined system O+P, which has eigenvalue 1 on the space spanned by all $|B,i\rangle \times |B,i\rangle$, and which has value 0 on all $|B,i\rangle \times |B,j\rangle$, for j not equal to i. Then if O and P are in pure dynamic states $|B,i\rangle$ and $|B,j\rangle$ respectively (always ignoring phase factors), the value of M will be 1 if and only if $i = j$. In the usual interpretation, this means that in the only case in which our pointer readings can have definite values, M will have the value 1 just in case these values agree. In the context of that interpretation, then, it is reasonable to speak of M as the observable which is, or registers, agreement between the two pointer readings. Even in the context of Rovelli's interpretation, one can continue to speak of M as the 'agreement' observable. The question is whether here, the locution needs to be taken with a grain of salt -may M take up the value 1 even though the pointer readings do not agree?

Our present example provides an illustration. The final dynamic state of O+P is $\sum |\beta_i|^2 P[|B,i\rangle \times |B,i\rangle]$. Since all summands of the mixture are eigenstates of M with eigenvalue 1, so is the state itself. So M takes the value 1 on the system O+P. To arrive at this conclusion however, we needed only to know the mixed state here ascribed to O+P -we did not need any information about what states O, P are in individually. That

information is logically compatible with the equally valid conclusion that O, P are in mixtures of the various states $\{|B, i\rangle\}$. So the conclusion that M takes value 1 cannot possibly, by itself, guarantee the suggestion that the pointer reading on P is equal to the pointer reading on O . But if we assume von Neumann's rather than Rovelli's interpretation, we do have that guarantee, since O, P collapsed into definite pointer states.

Reply: Once again, we have drawn a puzzling consequence for Rovelli by thinking about the situation in 'old' terms, and then having too quick a look at how his view differs. To really see whether there is a puzzle here, we have to retell the story from the beginning, in Rovelli's way. **Here is the retelling**, which we can now exhibit as a more elaborate example of Rovelli's view:

EXAMPLE 2: Enter third observer, ROV

We describe the situation from the point of view of a third observer, ROV. He has made measurements on S, O , and P in the past. On this basis he can say that the initial state of measured system S is a superposition $\sum \beta_i |A, i\rangle$ of eigenstates of an observable A corresponding to distinct eigenvalues, and that A will be measured twice, by two observers (measuring systems) O and P ¹⁸. The pointer observable on both O and P is B , with eigenstates $\{|B, i\rangle\}$.

Then, just calculating the time evolution on that basis, the combined system $S + O$ will be in dynamic state $\sum \beta_i (|A, i\rangle \otimes |B, i\rangle)$ relative to ROV at the end of the first measurement. At the end of the second measurement the dynamic state of $S+O+P$ relative to ROV will be $\sum \beta_i (|A, i\rangle \otimes |B, i\rangle \otimes |B, i\rangle)$.

By reduction, parts of the total system also have states relative to ROV at those times. As before, write $P[|A, i\rangle]$ for the projection on the ray containing $|A, i\rangle$, etc. At the end of the first measurement, the individual systems S and O are in states $[\sum |\beta_i|^2 P[|A, i\rangle]]$ and $\sum |\beta_i|^2 P[|B, i\rangle]$ relative to ROV, respectively. At the end of the second measurement, the final state of $O+P$ relative to ROV, also calculated by reduction, is $\sum |\beta_i|^2 P[|B, i\rangle \otimes |B, i\rangle]$.

¹⁸ It is part of ROV's knowledge, based on past measurements, that I and O each evolve freely after their measurement interaction, that there is no interaction between O and P , and that both A and the 'pointer-reading observable' B for O commute with the free Hamiltonians for S and O respectively.

There is *no need* to carry out a **third** measurement, of the 'agreement observable' M , because it is predictable with certainty by ROV that he will get 1 if he does.

But suppose now that ROV asks himself what O and P found, and whether they found the same thing. Then he is asking a question that has no answer, for he cannot answer questions about their past given that he made no measurements on the basis of which he could answer those questions!

Now, of course, ROV can decide to make two new separate measurements on O and P, to see what they are registering now. So suppose he measures $I \otimes B$ on O+P and gets value k . At this point he can make a prediction with certainty of what he will find if he then measures $B \otimes I$ on this system: for now the state of O+P relative to ROV is the result of conditionalizing the one he had, on this result. He predicts with certainty that he will see the same pointer reading $|B,k\rangle$ on O.

Was k the value that O and P saw at that earlier time? At this point we have no basis for thinking that this question can make sense on Rovelli's view. There are no states of O, P relative to ROV which could be consulted to answer it.

So, to summarize: with a von Neumann mindset we insist that there must be a fact of the matter about what O and P saw, *tout court*, and that a fact of the matter is always enshrined in a definite quantum state. But in Rovelli's world that is not the way things are.

5. CAN WE GO BEYOND THE RESOLUTION OF THESE PUZZLES?

What we have seen is that the puzzles one might have at first sight of Rovelli's account can be resolved. But the resolution leaves one still uneasy, for it hinges on the point that an observer O can register a measurement outcome –e.g. the answer 1 to question? A– but this fact is not equivalent to O being in a particular physical state, whether relative to itself or relative to any other observer.

In other words there are elements of Rovelli's 'meta' description which may in particular cases not correspond to any information had by

any observer, and hence *apparently* not describable in the language of quantum mechanics. One might be tempted to introduce the fiction that there is a 'universal observer' who knows what information is had (what answers have been registered) by each 'ordinary' observer. But this fiction can certainly not be admitted without ruining the story.

At the same time, in our reflections on what the observers register as measurement outcomes, we are targeting the very basis of Rovelli's understanding of quantum mechanics, and the very basis of the description of Rovelli's world:

Quantum mechanics is a theory about the physical description of physical systems relative to other systems, and this is a complete description of the world (Sect. IIC).

Drawing on Rovelli's favorite illustration of different frames of reference in Einstein's world, we are clearly tempted to ask: but *what relations are there between the descriptions that different observers give when they observe the same system?* Of course there can be no clue at all to an answer if we assume that there are no interactions at all between these distinct observers. But perhaps we can get a clue if we think of those distinct observers as themselves subject to observation by a third observer!

Doing so need not be illegitimate if we recall that Rovelli is describing the general form that any ascription of states or observable-values can take, and that this is the form of information that an observer *could* have.

5.1 Rovelli's symbolism for the information held, simplified

Let us take a look back at how, in his 'meta' description, Rovelli introduces a symbolism to express the fact that a given system 'has' information about another one:

If there is a maximal amount of information that can be extracted from the system, we may assume that one can select in $W(S)$ an ensemble of N questions Q_i , which we denote as $c = \{Q_i, i=1, N\}$, that are independent from each other. There is nothing canonical in this choice, so there may be many distinct families c, b, d, \dots of N independent questions in $W(S)$. If a system O asks the N questions in the family c to a system S , then the answers obtained can be represented as a string that we denote as

$$sc = [e_1, \dots, e_N]c \quad (4)$$

The string sc represents the information that O has about S , as a result of the interaction that allowed it to ask the questions in c (Section III-C).

The idea of a state of S relative to O enters now, because on the basis of this information, O can locate S in a finite subspace of the pertinent Hilbert space –even assign it a particular pure state represented by a vector in that space if the question-answer sequence was a maximally compatible one. This is what we describe informally in:

(Form 1) O registers answer 1 to complete question? A , so S has state $|A, 1\rangle$ relative to O .

We observe now that there is in effect a **time order**: the order in which the questions are asked. (Only order in time will be regarded for now, not time metric.) The N questions in numbered line (4) appear in the order $1, \dots, N$ so we can think of them as time-points, and can suggestively take them to indicate times t_1, \dots, t_N . But then the less formal description of (Form 1) should be expanded to the form:

(Form 2) O registers answer e_1 to complete question? $A(1)$ at time t_1 , so S has state $|A(1), e_1\rangle$ relative to O at t_1, \dots, O registers answer e_N to complete question ? $A(N)$, so S has state $|A(N), e_N\rangle$ relative to O at time t_N .

Moreover, in general O will calculate evolved states for periods between measurements (compare the formulations by Groenewold where this is made explicit). So I suggest that we can therefore speak of an evolving relative state, as follows:

(Form 3) S has state $|\psi(t)\rangle$ relative to O , during the interval (t_1, t_N)

or, when we note only certain special moments in that interval, the following is an acceptable form of description:

(Form 3-FIN) S has states $|\psi(1)\rangle, \dots, |\psi(N)\rangle$ relative to O , at times t_1, \dots, t_N .

In some contexts it will be convenient to suppress the time reference, and just use Form 1, but in other contexts we will have to use the full form 3 or 3-FIN.

5.2 Concrete example retold as by ROV

As concrete example let us take the situation in EXAMPLE 2, introduced in Puzzle 3, in which observers O and P were themselves subject to observation by an outside observer ROV, who started with the same initial information about S+O that P had, but also information about P, so that he can foresee the sequence of two measurements that were displayed in the example.

I will now designate the initial time as t_0 and the ending times of the two measurement interactions as t_1 and t_2 .

(States relative to ROV)

- a) The measured system S starts off in a state $\theta = \sum \beta_i |A, i\rangle$ relative to ROV, which is a superposition of eigenstates of an observable A corresponding to distinct eigenvalues.
- b) A will be measured twice, by two measuring systems O and P. Each of O and P will be in the 'ready to measure' state relative to ROV to start, $|B, r\rangle$. The indicator states are $|B, i\rangle$ for eigenvalues i of A (which do not include r).

We assume that I and O each evolve freely after their measurement interaction, that there is no interaction between O and P, and that both A and the 'pointer-reading observable' B for O commute with the free Hamiltonians for S and O respectively.

- c) The combined system S+O will be in state $\sum \beta_i (|A, i\rangle \otimes |B, i\rangle)$ relative to ROV at t_1 , the end of the first measurement.
- d) Similarly at that time, taking into account the as yet unchanging P, the state of S+O+P relative to ROV will be the superposition $\sum \beta_i (|A, i\rangle \otimes |B, i\rangle \otimes |B, r\rangle)$.
- e) At t_2 , the end of the second measurement the state of S+O+P relative to ROV will be $\sum \beta_i (|A, i\rangle \otimes |B, i\rangle \otimes |B, i\rangle)$.

Further states relative to ROV

When we look at d) and e) above, we see that the state of S relative to ROV does not change after t_1 , because the coefficients in the superposition do not change, even though the components do.

To show this, note that by reduction, we have states also for parts of the total system, namely S, O, P, relative to ROV. As before we write P [$|A, i\rangle$] for the projection on the ray containing $|A, i\rangle$, etc. We deduce

f) ROV assigns to S all by itself an evolving *mixture* $\rho(S, ROV)(t)$ of the states $|\psi(t, i)\rangle$ such that:

$$\begin{aligned} \text{for } t < t_1 \text{ the state } |\psi(t, i)\rangle &= \sum \beta_i |A, i\rangle, \\ \text{for } t_1 \leq t \text{ the state } |\psi(t, i)\rangle &= |A, i\rangle \end{aligned}$$

This mixture has as components the projections on these evolving vectors, one for each value i such that the coefficient β_i is not zero, and the weights are the 'squares' of those coefficients.

Note well, that *there is no change in this relative state at the second measurement time*, since in the superposition for the entire system, the values of B in O and P are the same in each component (that is, for every eigenvalue i such that β_i is not zero) from that moment on. Hence the weights in the mixture do not change from t_1 on.

5.3 ROV observes five measurements

In fact, by the definition of von Neumann type measurements –entirely in terms of the quantum mechanical states and evolution operators (Hamiltonians)– there are five such measurements in the situation we described just now! The initially given measurements are:

- a measurement of A by O ending at intermediate time t_1
- a measurement of A by P ending at final time t_2 .

Both of these have pointer observable B, and the criterion they meet, to count as von Neumann measurements, is that the interaction is such that

(vN Criterion) beginning state $|A, k\rangle \otimes |B, r\rangle$ of system S+O evolves into $|A, k\rangle \otimes |B, k\rangle$, where k is any eigenvalue

of A; and a fortiori, beginning state $(\Sigma\beta_i | A, i\rangle \otimes | B, r\rangle$
 evolves into $(\Sigma\beta_i | A, i\rangle \otimes | B, i\rangle)$

and similarly for P. By the same token, there is also:

a measurement of A by O, also ending at t_2

where for simplicity we take A not to be time-dependent (if it is, the same holds, but the state of S relative to O evolves, in a way that O can calculate, and so adjust with time passing –no need, for our argument, to cover the general case). So O just keeps showing a value for A, and assigning the corresponding eigenstate to S, for all times from t_1 on.

But there is more. From the above it follows that the interaction between S and the total system O+P is also the correlate of a measurement –in fact of three distinguishable measurements. For example, if we take $B \otimes I$ and $I \otimes B$ respectively as pointer observables on O+P, then the vN criterion is satisfied for times t_1 and t_2 respectively. So we have:

two measurements of A by O+P, ending at t_1 and t_2 respectively.

To see this, in the story as told in terms of states relative to ROV, let us look at the overall evolution of the system, relative to ROV. At the final time t_2 , the complete system S+O+P is in pure state $\Sigma\beta_i(| A, i\rangle \otimes | B, i\rangle \otimes | B, i\rangle)$ relative to ROV. By reduction the other states relative to ROV are:

S is in $\Sigma |\beta_i|^2 P[| A, i\rangle]$
 O and P are both in $\Sigma |\beta_i|^2 P[| B, i\rangle]$
 S + O and S + P are both in $\Sigma |\beta_i|^2 P[| A, i\rangle \otimes | B, i\rangle]$
 O+P is in $\Sigma |\beta_i|^2 P[| B, i\rangle \otimes | B, i\rangle]$

Inspection shows that the vN Criterion is satisfied for the interactions I mentioned. But we can add one more: taking $B \otimes B$ as pointer observable, we also see O+P engaged in a measurement that ends at the later time t_2 . That is the fifth measurement which appears in this story of the states of these various systems relative to ROV, and their various evolutions¹⁹.

¹⁹ There can be no objection, it seems to me, to allow for trivial limiting cases: if O has absolutely no interactions with S through which information is gained, it is only a matter of bookkeeping if we say that then the state of S relative to O is the represented by the Identity operator –the ‘informationless’ statistical operator. This convention may at times smoothen the presentation, even if it is not really needed.

The reason it is important to note this is of course that observers gain information about systems only by measurement, and it is only if they gain information about systems that those systems have states relative to them. So now we can continue, in accordance with the meta-description of Rovelli's world, to see what states S has relative to O , P , and $O+P$.

States relative to O , P , $O+P$

We can find the states of S relative to O , to P , and to $O+P$ for that interval, except that there will be some unknowns in it, namely the eigenstates that these observers assign to S on the basis of the measurements they make on it. (ROV makes no measurements on S , after the interval begins, that is why there are no similar unknowns in our calculation of $\rho(S, ROV)$). So we arrive at:

- (g) Observer O assigns to S an evolving pure state $\rho(S, O)(t)$:
 for $t < t_1$ the state $\rho(S, O)(t) = \theta = \sum \beta_i |A, i\rangle$,
 for $t_1 \leq t$ the state $\rho(S, O)(t) = |A, m\rangle$

and here the value m is an unknown, it is the result that O registers as outcome of the measurement.

- (h) For P it is only a little more complicated: P assigns to S a mixture $\rho(S, P)(t)$ of the evolving pure states $\lambda(t, i)$ with weights $|\beta_i|^2$:
 for $t < t_1$ the state $\lambda(t, i) = \theta = \sum \beta_i |A, i\rangle$,
 for $t_1 \leq t < t_2$ the state $\lambda(t, i) = |A, i\rangle$,
 for $t_2 \leq t$ the state $\lambda(t, i) = |A, k\rangle$

and here the value k , the outcome of P 's measurement is unknown.

In the case of $O+P$ we see that it is an observer who makes two measurements, one precisely at the time of O 's measurement, and one at the time of P 's measurement, and finds respectively at that time the values r and s -two unknowns for us, as for ROV, for we have no basis or law on which to connect the outcomes of measurements by different observers, no matter how intimately they may be related. But just as did O , this observer does not assign a mixture, it assigns the pure state

$$\begin{aligned} \text{(i) } \rho(S, O+P)(t) &= \text{the state } \theta = \sum \beta_i |A, i\rangle, \text{ for } t < t_1 \\ &= \text{the state } |A, r\rangle, \text{ for } t_1 \leq t < t_2 \\ &= \text{the state } |A, s\rangle, \text{ for } t_2 \leq t \end{aligned}$$

with r and s as the unknowns.

We would like to see what constraints could be added that would ensure+ concordance between the states of a system S relative to different observers such as O , P , and ROV –and here it will be pertinent for us that we have to keep also $O+P$ in view.

So now, finally, I'm going to propose an addition to Rovelli's account.

5. 4 Additional Postulate relating relative states

Additional Postulate For any systems S , O , P , witnessed by ROV :

- the state of S relative to O (if any) cannot at any time be orthogonal to the state of S relative to $O+P$ (if any), and
- the state of S relative to P (if any) similarly cannot be orthogonal to the state of S relative to $O+P$ (if any),
- and the state of S relative to any of these cannot be orthogonal to the state of S relative to ROV ,
- (and so forth for larger composite situations).

Here too the words “if any” are needed for generality; but in our example, the three systems do assign states to S . We may note again that the case of pure states is very special, and in general (as opposed in our examples here) the relational states will be mixed –and there is no associated ‘ignorance interpretation’ of mixtures. The requirement of non-orthogonality is rather restrictive for pure states, but of course always less so for mixtures.

What could be the motivation and intuitive warrant for this postulate, within the point of view of relational quantum mechanics? As Rovelli presented his own motivation he refers to the example of Einstein's methodology in the creation of relativity theory in just the same way that the Copenhagen physicists took their inspiration from that episode. The inspiration took the form of a certain kind of moderate empiricism: nothing was to be attributed to how nature itself is or proceeds beyond what is manifested in measurement outcomes. Thus the overriding case

for the denial that certain observables really do have simultaneous sharp values when not measured is precisely that there is no measurement procedure to reveal that possibility. More precisely, no configuration of values of observables is to be postulated for unmeasured nature unless there is a state in which measurement would show that configuration as outcome, with certainty.

In the quantum case, where transition probabilities are zero precisely when the relevant states are orthogonal, we can encapsulate this idea in the

Slogan: Born probability =0 \rightarrow NO!

So consider how the situation looks to ROV. When ROV contemplates measurements on these systems, to see if the pointer observables of O , P , $O+P$ could be in disagreement with each other at the pertinent times, the calculation of the Born conditional probability for this will be zero. So, to follow the above suggestion as to how to conceive of the unmeasured world, ROV will conceive of the relations between what the subsystems register accordingly. The idea that any assertion about what happens in nature must have cash value in what we can expect to detect, measure, or observe is strong in the Copenhagen tradition, even if contradicted by hidden variable enthusiasts. It seems to me that it echoes precisely the sort of inspiration that both the Copenhagen theorists and Rovelli derive from Einstein's reasoning when he introduced relativity.

So how is this inspiration honored by our Additional Postulate? If we now look back to our description of the evolving states of S , through the relevant time interval, relative to these three observers, we see the following pure state assignments:

$\rho(S, O)(t)$ remains the same from t_1 on, namely $|A, m\rangle$
 $\rho(S, P)(t)$ is a mixture until t_2 when it becomes $|A, k\rangle$
 $\rho(S, O+P)(t)$ is $|A, r\rangle$ for times from t_1 on, till it becomes $|A, s\rangle$ at t_2

For different values of m, k, r, s , those vectors are mutually orthogonal, since they are all eigenvectors of the same operator. So the second and third line immediately tell us that $k = s$. But the first and third line tell us that $m = r$ when we attend to t_1 , and similarly that $m = s$, when we attend to t_2 . So all these numbers are after all the same.

Result: the evolving states of S relative to the observers O and P are not the same to begin, but they are the same once P makes its A-measurement on S, sometime after O did (with no disturbance of A intervening meanwhile).

Supposing ROV to be knowledgeable of Relational Quantum Mechanics *thus extended*, what can he know even though he has made no measurements during or after that interval, on O and P?

He knows that what they found as outcomes of their measurements were indeed the same.

He already knew on the basis of Quantum Mechanics alone that *if he made a measurement to check on such agreement* he would get the answer YES with certainty. But now, calculating from the same previous measurement results that constitute his initial information, but using also Additional Postulate, he deduces that the agreement he would find with certainty if he measured was indeed already there.

This pleasing result, I have to emphasize, is found only by adding this additional postulate concerning how *the information registered* by components of a composite system engaged in several measurements are related to each other. So I cannot pretend that this harmony between the information obtained by different interacting observers follows from what Rovelli presented.

But Rovelli did not go into the question of whether there are three-way connections between information that can be had by observers in such a larger situation. This additional postulate was phrased so as to add only to the general form in which information can be had by different systems in a complex situation –without ever assigning any quantum mechanical states that are observer-independent.

I submit that the addition is consistent with Rovelli's account, and does not go essentially beyond what Rovelli allows himself in the 'meta' description in which he couches his depiction of the world of quantum mechanics. For it remains that all that has been provided – once we recognize the holism in composite situations involving many interactions– is an answer to "what is the general form of a description of the world from the vantage points of different observers?"

6. RELATIONAL EPR

Laudisa 2001 and Smerlak and Rovelli 2006 have examined how the Einstein-Podolski-Rosen situation can be regarded or modeled within Relational Quantum Mechanics. They do not entirely agree in their approach. Here I shall show how the situation fares if my Additional Postulate is accepted. The result appears to be different from what is favored by Rovelli, though it does not seem to affect the empirical content of the resulting formulation of quantum mechanics.

Let S be a two-part system $\alpha+\beta$ (such as a photon pair in singlet state), in a superposition of correlated states $\uparrow\otimes\downarrow$ and $\downarrow\otimes\uparrow$. The arrows are eigenvalues of observable A .

Observers $P1$ and $P2$ respectively measure $A\otimes I$ and $I\otimes A$ with pointer observable B . ROV has information on initial states and dynamic process

$P1$ gets \uparrow or \downarrow ... the state of α relative to $P1$ is $|\uparrow\rangle$ or $|\downarrow\rangle$
 $P2$ gets \uparrow or \downarrow ... the state of α relative to $P2$ is $|\uparrow\rangle$ or $|\downarrow\rangle$
 $P1+P2$ gets $\uparrow\uparrow$ or $\uparrow\downarrow$ or $\downarrow\downarrow$ or $\downarrow\uparrow$... the state of $\alpha+\beta$ relative to $P1+P2$ is $|\downarrow\rangle\otimes|\uparrow\rangle$ or $(|\uparrow\rangle\otimes|\uparrow\rangle$ or $(|\downarrow\rangle\otimes|\downarrow\rangle$ or $(|\uparrow\rangle\otimes|\downarrow\rangle$

For $P1+\alpha+P2+\beta$ ROV assigns at the measurement time a superposition of

$$(|B, 1\rangle\otimes|\uparrow\rangle)\otimes|B, 2\rangle\otimes|\downarrow\rangle$$

and:

$$(|B, 2\rangle\otimes|\downarrow\rangle)\otimes(|B, 1\rangle\otimes|\uparrow\rangle)$$

This implies that ROV assigns to $\alpha+\beta$ a mixture of $(|\uparrow\rangle\otimes|\downarrow\rangle)$ and $(|\downarrow\rangle\otimes|\uparrow\rangle)$.

By the Additional Postulate it follows that the state of $\alpha+\beta$ relative to $P1+P2$ must be one of these, thus ruling out two of the possibilities noted above. And then $P1+P2$ will assign to α and β separately either $|\uparrow\rangle$ and $|\downarrow\rangle$ respectively or $|\downarrow\rangle$ and $|\uparrow\rangle$ respectively. But then, again by the Postulate, the states of α and β relative to $P1$ and to $P2$ respectively cannot be the same, on pain of orthogonality to what they are relative to $P1+P2$.

Have we arrived at ‘spooky’ non-locality? We need to be worried by possible conflict with the sentiment so clearly expressed in Smerlak & Rovelli:

There is no operational definition of observer-independent comparison... of different observers’ information (...): the information of different observers can be compared only by a physical exchange of information between the observers²⁰.

Can ROV, in our story (including the Additional Postulate), compare the two states of $\alpha+\beta$ relative to P1 and P2 before measuring P1 and P2 at the end?

YES and NO!

ROV can know that P1 and P2 did not register the same value for A. But to know anything about *which* values they did register, ROV would have to make measurements. So ROV can *predict* no more than someone who has not heard the additional Postulate, but only that *if he measured both* he would find different registered values, which is predictable with no reliance on the additional Postulate.

We can define a function of the outcomes registered by P1 and P2, which takes value 1 if the outcomes are the same and value 0 if they are different. It seems then that ROV can know the value of this defined quantity, without having measured it. But the defined quantity has value 1 if and only if the states of α and β relative to P1 and P2 respectively are either $|\uparrow\rangle$ and $|\downarrow\rangle$, or $|\downarrow\rangle$ and $|\uparrow\rangle$. That this is so I do not think follows in the original Relational Quantum Mechanics. Therefore this could be counted as running contrary to the above cited sentiment.

But I would like to suggest that it may count as a reason for the suggested Additional Postulate. For otherwise we leave open the possibility that the state of $\alpha+\beta$ relative to P1+P2 is $(|\uparrow\rangle\otimes|\downarrow\rangle)$, for example, although the states of α and β relative to P1 and P2 respectively are both $|\uparrow\rangle$!

Even if we were to insist that “S has state ... relative to O” can only have a truth value related to a further observer ROV (and not be true

²⁰ SMERLAK, M. & ROVELLI, C. Relational EPR. 1996. ms.

or false 'absolutely') this same difficulty would appear when ROV is in the picture²¹.

But there is much to explore here yet, including the most radical view, namely that even what the states relative to any observer are must itself be relative to an observer.

Appendix. An auxiliary formalism

Finally, *solely as an aid to the imagination*, we can add some auxiliary symbolism, as follows. We note that O registers an answer in an entirely physical way, in that this measurement involves –and requires– a correlation of the measured observable A on S with a 'pointer' observable B on O. Hence, if we wish to mark that correlation, we have a final variant on Form 1:

(Form 1-bis) O has $[|B, 1\rangle]$

This looks deceptively like a state attribution to O, which it is not –it is meant as an equivalent to Form 1, when it is known that O's pointer observable is B, and no more.

We can think of this as an encoding of the information O has about S in the spirit of Rovelli's remark:

Let me then take a lexical move. I will from now on express the fact that q has a certain value with respect to O by saying: O has the "information" that $q=1$ (Section II-E).

In this passage, q is an observable that pertains to a specific system, the one that is measured by O, so despite the surface form this is still a relational statement. A specific example will have a projection on a subspace for q, and a still more specific example will have this projection operator one-dimensional, in which case to say ' $q=1$ ' is the same as ascribing a specific pure state (to the measured system S, relative to O).

²¹ The suggestion here would be, it seems to me, a radicalization of the original Relational Quantum Mechanics, but perhaps closer to the initial intuitions than what I have worked with here. However, though worth exploring further, I see it as difficult to sustain, given its obvious danger of either regress or circularity –well, perhaps worth exploring precisely because of such danger!

In the October 2006 symposium at the University of Provence, Aix, Carlo Rovelli voiced some suspicion of this auxiliary symbolism, so I undertook to restate the argument without reliance on that device.

Using the auxiliary symbolism, however, we have an equivalent alternative to the Additional Postulate:

If a composite observer $X+Y$ has $[[\varphi]]$, while X has $[[\xi]]$ and Y has $[[\xi']]$ then ξ is possible relative to reduced state $\# \varphi$ and ξ' is possible relative to reduced state $\varphi\#$.

Written in this way, one can see a formal relationship –though well short of agreement throughout– with the modal interpretation (CVMI) as defended in my 1997. There, and in the there cited section of my 1991, it is possible to see just how the last line in the Additional Postulate (“and so forth for larger composite situations”) would need to be elaborated in detail.

REFERENCES

- BUB, J. (2004). “Why the quantum?” *In*: Studies in the history and philosophy of modern physics, vol. 35, p. 241-266.
- _____. (2005). “Quantum Mechanics is about quantum information”. *In*: Foundations of Physics, vol. 35, p. 541-560.
- DICKE, R. H. (1989). “Quantum measurements, sequential and latent”. *In*: Foundations of Physics, vol. 19, p. 385-395.
- FUCHS, C. A. (1957). “Objective and subjective aspects of statistics in quantum description”. *In*: KÖRNER, S. (Ed.). Observation and interpretation in the philosophy of physics proceedings of the ninth symposium of the Colston Research Society, p. 197-203.
- _____. (1998). “Information gain vs. state disturbance in quantum theory”. *In*: Fortschritte der Physik, vol. 46, no. 4-5, p. 535-565. Reprinted in BRAUNSTEIN, S. L. (Ed). Quantum Computation: where do we want to go tomorrow? (pp.229-259). Weinheim: Wiley-VCH Verlag, 1999.
- _____. (2004). “Quantum Mechanics as Quantum Information (and only a little more)”. *Quant-ph/0205039*.
- GROENEWOLD, H. J. (1952). “Information in Quantum measurement”. *In*: Koninklijke Nederlandse Akademie van Wetenschappen, B55, p. 219-227.

HAGAR, A. & HEMMO, M. (2006). "Explaining the unobserved—why QM isn't only about information". In: *Foundation of Physics*, vol. 36, p. 1295-1324.

LAUDISA, F. (2001). "The EPR argument in a relational interpretation of Quantum Mechanics". In: *Foundation of Physics Lett*, vol. 14, p. 119-132.

LAUDISA, F. & ROVELLI, C. (2005). "Relational Quantum Mechanics". In: *The Stanford Encyclopedia of Philosophy*. [On line] <http://plato.stanford.edu/entries/qmrelational/>

ROVELLI, C. (1996). "Relational Quantum Mechanics". In: *International Journal of Theoretical Physics*, vol. 35, p. 1637-1678.

SMERLAK, M. & ROVELLI, C. (2006). *Relational EPR*. ms.

TEMPLE, G. (1948). *The general principles of Quantum Theory*. London: Methuen.

Van FRAASSEN, B. C. (1991). *Quantum Mechanics: An Empiricist View*. Oxford: Oxford University Press.

_____. (1997). "Modal interpretation of repeated measurement: Reply to Leeds and Healey". In: *Philosophy of Science*, vol. 64, p. 669-676.

WHEELER, J. (1957). "Assessment of Everett's 'Relative State' formulation of quantum theory". In: *Review Mod. Phys.*, vol. 29, p. 463-465.

A LITTLE GIVE AND TAKE: PROBLEMS IN THE EMPIRICISM OF SELLARS AND HIS FOLLOWERS

UN POCO DE TOMA Y DADA: PROBLEMAS EN EL EMPIRISMO
DE SELLARS Y SUS SEGUIDORES

MICHAEL DAVIS

University Australian National, Australia.
Universidad de los Andes, Colombia. discufilo@ucaldas.edu.co

RECIBIDO EL 15 DE SEPTIEMBRE DE 2010 Y APROBADO EL 30 DE SEPTIEMBRE DE 2010

RESUMEN ABSTRACT

El punto de inicio de este artículo es el rechazo de parte de Sellars del empirismo fundacionalista como se encuentra en su discusión del *mito de lo dado*. Sellars ataca el mito desde dos ángulos principales que corresponden a los dos elementos del empirismo: la idea que nuestras creencias son justificadas por el mundo, y la idea que nuestros conceptos se derivan de la experiencia. Correctamente atacando la segunda, Sellars también está incorrectamente atacando la primera. Por ende, Sellars rechaza la idea del sentido común que, por lo menos algunas de nuestras ideas se justifican por apelación al mundo empírico. Mi propósito es descubrir las suposiciones que llevan a Sellars a este punto, y como operan las mismas suposiciones en sus seguidores, como Brandom, Rorty y McDowell. Y luego mostrar como un rechazo de dichas suposiciones nos dan una manera de superar este problema al no recaer en el fundacionalismo.

The starting point of this paper is Sellars's rejection of foundationalist empiricism as found in his discussion of the *Myth of the Given*. Sellars attacks the Myth from two main angles, corresponding to the two elements of empiricism: the idea that our beliefs are justified by the world, and the idea that our concepts are derived from experience. In correctly attacking the second, Sellars is also, incorrectly, led to attack the first. Thus, Sellars rejects the commonsensical idea that at least some of our ideas can be justified by appeal to the empirical world. My purpose is to examine why Sellars is led to this point, and how the same assumptions that lead him there also operate in his followers, such as Brandom, Rorty and McDowell. I then show how a rejection of these assumptions gives us a way around this problem that does not fall back into foundationalism.

PALABRAS CLAVE KEY WORDS

contenido conceptual, empirismo, fundacionalismo, McDowell, Mito de lo dado, Sellars.

conceptual content, empiricism, foundationalism, McDowell, Myth of the Given, Sellars.

1. SELLARS'S MYTH OF THE GIVEN

The *Myth of the Given* arises, according to Sellars, from a confusion of empirical explanation with justification, from an attempt to subsume what he calls the space of reasons under the space of causes. This, he says, is a kind of naturalistic fallacy, for in seeking a foundation for knowledge, we are seeking to justify what we know, to establish its epistemic value. This is an operation in the space of reasons, not in the space of causes. "In characterizing an episode or a state as that of knowing, we are not giving an empirical description of that episode or state; we are placing it in the logical space of reasons, of justifying and being able to justify what one says"¹.

Empiricism seeks to justify our conceptualization of reality by deriving it or reconstructing it from our sensory experience of the world. In this sense modern empiricism is a form of Aristotelian epistemology, minus the Aristotelian metaphysics which make it coherent. Sellars correctly rejects this approach, arguing that our conceptual scheme must be autonomous with respect to the world, that it must be an exercise in spontaneity in the Kantian sense. At the same time, however, or indeed by the same token, on Sellars's account the justification of our beliefs is independent of any semantic relations with the world, for the world exists in the space of causes and justification cannot transcend the space of reasons. Reality cannot act as a constraint on judgement. Justification is a matter of warranted assertibility operating within socially sanctioned inferential moves.

The problem here is that it seems intuitive that reality is and should be a constraint on human reason. Our ideas about reality -be they simple beliefs or complex conceptualizations- should in turn be judged against that reality. Sellars establishes the freedom of human reason to conceptualize and to judge as it sees fit, but this freedom is empty since it cannot be responsible to reality. In what follows I seek to uncover the assumptions that create this problem, and to show that a rejection of the Given need not imply a rejection of these intuitions about rational constraint.

Sellars's attack on the Given is multifold. Let us have a look at several of the different strands of his argument.

¹ SELLARS, W. *Empiricism and the Philosophy of Mind*. Cambridge: Harvard University Press, 1997.

Sensory episodes cannot justify belief

Sellars rejects the idea that our perceptual beliefs may be justified by what he calls 'indubitable episodes' of sensing. These episodes are held to be indubitable because they are the end points in a causal chain that begins in the world. They have no conceptual form, but are thought to justify beliefs, which do have conceptual form. This, according to Sellars, is illegitimate: the non-conceptual cannot justify the conceptual.

Sensory experience presupposes a conceptual apparatus

Empiricism attempts to justify our very conceptualization of the world by deriving it from these same sensory episodes, or by explaining the origin of our concepts in such episodes. Sellars argues that perceptual knowledge, insofar as we can call it knowledge -that is insofar as it can participate in the justification of knowledge claims in the space of reasons- presupposes the possession of an acquired conceptual framework. We have no epistemic relation with experience that is not at the same time a conceptual relation. Observation reports, even if they in some sense justify empirical propositions, themselves require a whole raft of beliefs and knowledge to be meaningful as observation reports. Moreover, knowledge, even perceptual knowledge, is a matter of judgement (of 'linguistic awareness'), and judgements essentially involve general concepts. General concepts transcend any particular experience and thus cannot be derived from observational acquaintance with mere particulars. Therefore our conceptual framework cannot be derived from unconceptualized experience.

Justification as inference

The view that sensory impressions are epistemically irrelevant because they cannot participate in inferences. Sellars argues that justification is achieved through inference. Since only propositionally or conceptually structured items can participate in inferences, sense experience, being non-propositional, cannot play a justificatory role. It might be argued that if sense experience is propositional or conceptual in form, then it can participate in such inferences. This is true, but according to Sellars it is of no use to the foundationalist because we thereby begin from propositions or beliefs, that is, unanalyzed knowledge claims, and these presuppose the conceptual apparatus that the empiricist wishes to derive from non-conceptual experience. This is what Sellars means by

the view he calls psychological nominalism, the view that all awareness is linguistic awareness.

Now Sellars's attack on the Given is correct insofar as it rejects the idea that indubitable sensory episodes are the foundation of empirical knowledge. Sense data cannot provide an indubitable foundation for empirical knowledge because, contra (1), they are not indubitable and, more importantly, contra (2), our conceptual scheme is autonomous with respect to experience and indeed with respect to the world itself.

Yet as we have seen, the way Sellars constructs his attack on the Given leads him to reject the commonsensical idea that reality is a rational constraint on judgement. In an important way, he takes this to follow from claim (2), the claim that our conceptual scheme is autonomous. For Sellars, the autonomy of concepts implies that the world cannot act as a rational constraint on knowledge. In this he is mistaken, as I will attempt to show in Section 3.

Note for now that Sellars's arguments against (1) and (2) rest on the view he calls psychological nominalism, the view that all awareness is linguistic awareness. And that this view seems strongly correlated with the idea that there are no non-inferential forms of justification. For Sellars, since justification is a matter of inference, then experience, insofar as it is epistemically relevant, must be conceptual in form. Any non-conceptual modes of experience are irrelevant to justification.

Hence the dubitability or otherwise of sensory episodes is inessential to Sellars's argument. What is essential is the idea that unconceptualized experience cannot justify knowledge claims. I will argue in Section 5 that this idea is mistaken. We have a form of cognitive access to the world that is not at the same time conceptual, but this access is not indubitable and therefore cannot provide a certain foundation for empirical knowledge claims.

Moreover, with Sellars, I reject the empiricist attempt to derive our conceptual schemes from sensory experience. For important reasons, as we will see, our conceptual schemes must be autonomous with respect to the world.

2. THE CONSEQUENCES OF SELLARS'S REJECTION OF THE MYTH

As we saw, Sellars argues that whatever is non-conceptual in form can have no rational bearing on what is conceptual in form. The implication of this view is that there are no semantic or justificatory relations between propositions and matters of fact, or between word and world, for such a relation would cross the divide between the space of reasons and the space of causes.

Yet at the same time, Sellars is not willing entirely to abandon the idea that the world acts as some kind of rational constraint on our knowledge. This clearly presents him with a difficulty, since the world is not conceptual in form and therefore cannot, on his view, act as such a rational constraint. But if the world does not act as such a rational constraint, then our knowledge is essentially unanchored and need bear no relation to the world at all. This is the view that we find in Richard Rorty. So Sellars has a problem, namely that the world both must, and cannot, act as a rational constraint on knowledge.

Sellars attempts to resolve this difficulty through appeal to a two-level conception of truth. It is this which ties his attack on the Given to his larger project of reconciling the two images of man, the manifest and the scientific.

Sellars's two-tiered view of truth makes a distinction between justification, which belongs entirely to the space of reasons, and picturing, which belongs entirely to the space of causes. The picturing relation is therefore not a word-world semantic relation (which would be to succumb to the *Myth of the Given*) but a factual relation between two objects (which Sellars calls a semantic uniformity). The adequacy of the picturing relation admits of degrees and is subject to possible improvement. This then is Sellars's solution to the problem of reconciling the two spaces. Eventually we will achieve a reunion of the two spaces through the development of more-adequate theories of the picturing relation that will allow us to operate with an ever-more-accurate picturing relation itself. Thus in the limit, assertibility in the space of reasons will be shown to match the relations that obtain between the picture and the pictured in the space of causes, and the picturing relation will be shown to be a true reflection of the pictured by the picture.

I will not go more into this view here. But suffice it to say that it is an example of the naturalist view that ordinary discourse is a kind of theory or proto-theory which is revisable in the light of empirical discovery and which, in the end, represents a kind of naturalized Hegelianism in that it places the locus of ultimate explanation in some unknown future rather than in any ordinary criteria we currently possess. It is thus driven by essentially metaphysical considerations. Of course, Sellars himself considered his philosophy to be exactly a kind of naturalized Hegelianism².

And with good reason: adopting some kind of Hegelianism is the ultimate solution for any modern naturalist philosophy that begins from the premise that there is a gap between the manifest and the scientific understandings of the world, and that our semantic criteria must ultimately be derived from the latter understanding. Since our ordinary understanding has no access to these criteria then we can never be sure that our judgements are true except with reference to an unknown future state in which the two understandings are reconciled³.

What is important for our current purposes is where this view leaves the notion of justification, which Sellars construes essentially as warranted assertibility in the space of reasons. The problem Sellars faces is that judgement, as it operates in the space of reasons, cannot be constrained by the world. The only way the world can act as a rational constraint on judgement, in Sellars's view, is as the equivalent of a Peircean ideal truth. This means that our judgements, as we operate with them in the space of reasons, have only an ideal connection with the world; yet what we seem to require to do justice to them as empirical judgements is a real connection with the world in which they can be shown to be true of the world as it is, and potentially shown to be so at the moment of judgement. That is, if the world is to be a rational constraint on judgement, there can be no unbridgeable gap between the world and the judgement: the world must be able to participate in our judgements to the extent that those judgements can be said to be of, and truly of, their objects. So given his insistence on this gap, in Sellars's picture judgements lack objective significance. Because the question of whether our acts of judging are correct is independent of the correctness of the

² Cfr. *Ibid.*, s20.

³ Analytic philosophy thus has a strong tendency towards Hegelianism, despite its origins in a rejection of the same. Of course there are other options, such as radical skepticism or an abandonment of the idea of truth altogether, as in Rorty.

judgements themselves, the idea that we are making objective, empirical judgements loses all meaning. The space of reasons could be a game that bears no relation to objective reality at all.

3. THE KANTIAN PREMISE

I now want to move to consider what it is that drives Sellars to adopt this view.

As we saw in Section 1, empiricism contains two distinct elements: the first is the idea that our beliefs can be justified by reference to sensory experience. The second is that our system of knowledge, or our conceptual scheme, can itself be shown to be an abstraction from sensory experience. Sellars's problem derives from the fact that he takes a rejection of the second element to entail a rejection of the first element as well.

Concepts, Sellars argues, are autonomous with respect to real objects. That is, neither our language, nor our conceptual structure, nor our theoretically constructed pictures, are determined by what exists. Our conceptual scheme is a product of spontaneity, in the Kantian sense.

That is, it is an exercise in freedom. We are free to construct and to modify conceptual frameworks, and these frameworks are not meaningful in virtue of whether or not they refer to real objects. If our conceptual frameworks were determined by what exists, or by impressions received through the senses, then we could make no sense of the idea of scientific progress as the development of ever-more-adequate conceptualizations of reality. We require something like this view because the development of modern science showed us that the behavior of objects on the empirical level could be explained only by appeal to a structure that is not immediately apparent but is instead accessible only theoretically through the empirical behavior of the objects. Where for Aristotle science and empirical reality were seemingly in harmony, the picture of reality that is given by modern science is often quite at odds with our phenomenal experience.

Moreover, our concepts are logically autonomous with respect to their objects because concepts are general. They therefore transcend, and thus cannot be simply derived from, any set of particular experiences. They

transcend their actual extension, thus their content cannot be reduced to this extension. In Aristotle, the metaphysics of forms means that the general can be derived from experience of the particular (meaning that in Aristotle concepts are not autonomous).

But modern empiricism rejects Aristotle's metaphysics in favor of nominalism, and in seeking to derive concepts from experience of empirical objects is therefore confronted with the problem of deriving the general from a finite set of experiences of particulars. Sellars instead adopts the rationalist view that the content of concepts is constituted solely by inferential rules.

Now it is precisely this view, that our conceptual scheme is not an abstraction from sensory experience, which leads Sellars to argue that the world cannot act as a rational constraint on our judgements. But it does so only because, as Macbeth points out, he makes the following assumption: "(F) If judgement is constrained by reality, then empirical concepts can be derived from experience"⁴.

Now it should be clear that this view Sellars shares with the foundationalist empiricism he rejects. That is, foundationalists suppose that knowledge can be founded on sense experience both in the sense that our beliefs can be justified by reference to that experience and in the sense that our concepts correctly reflect the world as it is just in virtue of being derived directly from our sensory experience of that world.

Moreover, it is precisely because of this assumption that Sellars is led to separate the space of reasons from the space of causes. Against foundationalism, Sellars denies the consequent, and thus concludes that reality cannot rationally constrain judgement. It is interesting that Kant assumes the same premise, but thinks it more plausible to affirm the antecedent.

He thus accepts the implication that empirical concepts are derivable from sense experience; to avoid empiricism, he argues that the empirical world is constituted by the mind. Thus, as Macbeth puts it, Kant is able to hold that judging is "at once (...) a standing in the space of reasons and rationally constrained by what is; but it can be only by virtue of the fact that what is known (...) is only an appearance"⁵.

⁴ MACBETH, Danielle. "Empirical knowledge: Kantian themes and sellarsian variations". In: *Philosophical Studies*, 2000. no. 101, p. 113-142.

⁵ *Ibid.*, p. 113.

Moreover, on Kant's view, concepts are not sufficiently autonomous to permit the idea that we are free to modify our conceptual scheme in accordance with how adequate we take it to be as a means to accurately describe empirical reality. As we saw, we must be free to refashion the concepts in terms of which we make empirical judgements. But for Kant, our conceptual scheme is ultimately founded on the synthetic a priori principles of mathematics and natural science.

But like Kant's, Sellars's solution is also inadequate. Sellars is right to deny the consequent of (F): our empirical concepts cannot be simply abstracted or derived from sense experience, for the reasons previously given; but he is wrong to conclude that the antecedent is false: reality does rationally constrain judgement. To avoid Kant's and Sellars' errors, it is necessary to see that it is the implication in (F) that is false. Even though reality can rationally constrain judgement, our conceptual scheme is autonomous.

Macbeth argues that the problem lies in the fact that the implication in (F) conflates two distinctions: that between concepts and objects on the one hand, which we find in the consequent; and that between spontaneity and receptivity on the other, which we find in the antecedent. Given this conflation, the fact that concepts are autonomous with respect to objects leads Sellars to the conclusion that spontaneity is autonomous with respect to receptivity, or sense experience.

Macbeth's identifying this conflation shows us what is wrong with the implication in (F), but it does not explain why Sellars takes the implication to be true. I believe it can be explained in the following way. Sellars believes that judgement can be rationally constrained by reality only if reality or our experience of it is conceptual and can therefore participate in inferences. But our conceptual schemes can be fully autonomous only if reality is nonconceptual and therefore external to the space of reasons. If reality were internal to the space of reasons it would, in the manner of Kant, at least partly constitute the rules that determine the meaning of our empirical concepts. Hence, if our conceptual schemes are to be autonomous then judgement cannot be rationally constrained by our experience of the world. We can therefore see that Sellars's fundamental assumption is the view that rational constraint is a relation that can hold only between conceptually structured items.

Against Sellars, I believe that we can accept both that our conceptual scheme is not founded in acquaintance with objects and that the

application of concepts to objects in judgement *can* be justified with respect to our perception of how things stand with the objects in the world. In other words, we can accept that reality acts as a rational constraint on judgement while rejecting foundationalism. Both the creation and modification of conceptual schemes and the application of concepts in empirical judgement are exercises in spontaneity, and their content is not beholden to the world. But at the same time, both conceptual schemes and judgements can be measured for their adequacy against the world. This requires us to abandon Sellars's view that rational constraint is a relation that can hold only between conceptually structured items.

4. SELLARS'S FOLLOWERS

At this point it is useful to state that many of Sellars's followers fall into this same confusion. These include Brandom, Davidson, and Rorty. Indeed Rorty's philosophy is in many ways just Sellars's absent the notion that an ideal completed science provides the ultimate criterion of truth. I will not look more into the views of these thinkers here. Suffice it to say that they suffer from the same problem as Sellars in accounting for the role of the objective world in the justification of belief.

Another important follower of Sellars is McDowell. Macbeth sees in McDowell at least a partial resolution of this problem⁶. Thus McDowell accepts both Kant's view that spontaneity cannot be understood independently of receptivity and Sellars's view that concepts and objects are distinct. Spontaneity and receptivity are constitutively related in the sense that the capacity to have a satisfactory standing in the space of reasons is by the same token the capacity to be open to the world. This is why for Kant spontaneity is an exercise in responsible freedom. We are free to judge as we please, but in doing so we are responsible to the world as it is. The fact that spontaneity and receptivity are internally related in this way also shows that truth as a semantic relation is essential to justification and cannot exist apart from the space of reasons as Sellars supposes.

McDowell pictures this in Sellarsian terms as the idea that the space of reasons in a way extends over the space of causes. Through receptivity

⁶ See MACBETH, Danielle. "An Antinomy of Empirical Judgment: Brandom and McDowell". Forthcoming, available at: www.haverford.edu/philosophy/dmacbeth/publications/

we hold a passive relation to the world as we encounter it; the world in turn shapes the space of reasons by providing facts for use within that space. Yet McDowell continues to subscribe to the Sellarsian, rationalist view that justification is a matter of inference, or in other words, that what is not conceptually structured cannot act as a rational constraint on judgement. Thus, for McDowell, if reality is to act as a rational constraint on judgement, reality as it is presented in sense experience cannot be located outside the realm of the conceptual. This means that experience must be conceptual in form, and so too, as it is presented in sense experience, reality. Moreover, reality is something with respect to which in experience we are wholly passive. Experience is the actualization of our conceptual capacities by the world. And in this way, "experience enables the layout of reality itself to exert a rational influence on what a subject thinks"⁷.

There are several problems with this view. First, it does not seem to do justice to the spontaneity of our conceptual scheme or indeed of judgement itself. Thus McDowell states that "experiential uptake (...) supplies the content, the substance that thoughts would otherwise lack"⁸.

So on his view; thoughts have content only to the extent that they have experiential or empirical content. But this cannot be true. Thoughts that we entertain in our imagination; newly conceived scientific hypotheses; dreams. All these are thoughts with content, with meaning, that is not provided through experiential uptake. We can think in these ways precisely because the content of thought is not equivalent to empirical content. Moreover, the problem remains even if we take McDowell's position to apply only to empirical judgements. If objects actualize empirical concepts, then concepts cannot be truly independent of objects and we fall back into foundationalism.

McDowell thinks that if the world does not provide the content of thought in this way, then it cannot act as a rational constraint towards it. But the world acts as a rational constraint on thought not by providing it with content, but by providing an object of comparison to our thoughts. That is, we judge a thought to be empirically true when what it expresses about the world is in fact the case, as McDowell himself says⁹.

⁷ McDOWELL, J. *Mind and World*. Cambridge: Harvard University Press, 1996.

⁸ *Ibid.*, p. 4.

⁹ *Ibid.*, p. 27.

But in order to make this judgement, we must be able to consider the thought, and in an act that is independent of thinking the thought, compare it to the world itself. What this means is that we must have a kind of engagement with the world that is different from, and prior to, the kind of engagement that is involved in making a judgement. Otherwise we can only compare judgement to judgement, belief to belief, and we find ourselves enclosed in a linguistic circle outside of which lies the world itself. Only if experience and judgement are independent in this way can we make sense of how it is possible to revise our concepts in the light of what we come to know about the world. Experience therefore cannot be a passive actualization of our conceptual capacities.

Thus, in order for our conceptual scheme to be truly autonomous with respect to the world; that is, in order for us to be truly free to conceive of the world as we wish (*i.e.* falsely as well as truly); judgement of the sort just described must depend in part on an engagement with reality that is not conceptual or propositional in the sense in which Sellars, McDowell, and the other Sellarsians believe it is. Otherwise, we can choose only the road taken by Brandom and Rorty, which holds that the world acts only as the generator of causal input into a belief-relating mechanism; or the road taken by McDowell, which holds that concepts are dependent on empirical content. The first road means that the world cannot act as a rational constraint on judgement; the second that it can do so only if we give up the full sense of the spontaneity of thought and of the autonomy of our conceptual scheme (which in the end is no rejection of the Given at all, as Rorty argues of McDowell)¹⁰.

Macbeth puts it in this way: Brandom collapses the semantic dimension of thought into conceptual content; and while McDowell does not quite do the reverse, he believes that the content of a concept must in part be determined by its relation to the world. Now Macbeth proposes that the way out of this dilemma is to make use of Frege's distinction between sense and reference¹¹. We can thus distinguish between the sense of a thought, which is independent of any semantic relation with the world, from its reference, which provides just that semantic relation. In this way we can maintain the autonomy of concepts from objects while also maintaining the possibility of relations of truth between thoughts and the world.

¹⁰ Cfr. RORTY, R. "The Very Idea of Human Answerability to the World: John McDowell's version of Empiricism". In: *Truth and Progress*. Cambridge: Cambridge University Press, 1998. p. 138-52.

¹¹ MACBETH. "An Antinomy of Empirical Judgment: Brandom and McDowell". *Op. cit.*, p. 12ff.

It is clear that something like Frege's distinction is logically required to make sense of judgement, though to what extent we conceive it along Fregean lines is a question I will not discuss here. However, though necessary, it is not sufficient to solve our problem. In addition, we require an alternative to the empiricist phenomenology of perception and epistemology of judgement. Frege's distinction on its own provides no route by which we can actually make the judgement I described above, in which we can hold out a thought for comparison with the world itself.

5. RECONCEIVING EXPERIENCE

As I suggested earlier, I believe a large part of the problem to stem from the view that justification is always a matter of inference. This seems to me to be incapable of making sense of perceptual knowledge. On this view, for an act of perceptual belief to be justified it must be supported through inference to other beliefs which we know to be justified. But this cannot be the case. In quite simple terms, we generally take our perceptual beliefs to be justified when we can see that the world is just how we perceive it to be. And this is not itself a matter of inference. If justification were only a matter of inference, we could never exit the circle of justification, and our premises would never be justified. Justification cannot be a matter of pure inference since a valid inference does not tell us whether the proposition is true.

So perceptual knowledge must be non-inferential. It can, after it is conceptualized, participate in inferences, but it is not itself justified by means of inference. Conceptualizing experience in the manner of McDowell is of no help in the justification of perceptual knowledge, though it explains, unnecessarily, how perceptual beliefs can be used to justify other beliefs by inference.

Moreover, since our concepts are independent of objects, the world itself is not already conceptualized, and experience cannot be the mere actualization of concepts by objects. If it were, then we could make no sense of the possibility that our concepts could be inadequate representations of those objects. So perceptual knowledge is (at least some of the time) justified by reference to the non-conceptual, that is, by reference to the world. McDowell might object that this is to succumb to the *Myth of the Given*. But the *Myth of the Given* with respect to the justification of belief is not essentially the idea that something

conceptually structured can be justified by reference to something not conceptually structured, e.g. Sellars's non-verbal episodes, but those parts of our knowledge can be founded on some kind of indubitable awareness.¹² It is thought that because the conceptual cannot be indubitable, since it is an exercise in spontaneity, taking all awareness to be conceptual cannot cause us to succumb to the *Myth of the Given*. Yet acts of perception are not indubitable, whether or not one takes them to be conceptual. What leads us to thinking that they must be, is a picture of perception according to which what we perceive is the ultimate effect of a causal chain leading from the world to our minds. Thus, once again, an epistemological picture has been created on the basis of essentially metaphysical assumptions (naturalism).

To see what is wrong with this view, let us think again about how we are able to form theoretical conceptions about our own modes of perceiving the world. We can theoretically explain the way the world appears to us precisely because we can form conceptions of the world that are not beholden to our experiences -that is, precisely because our conceptual schemes are spontaneous-. And we can test these theories against the world itself, again precisely because we can perceive the world independently of our conceptual schemes. And what this in turn requires is a rejection of the model of perception according to which the world just causes beliefs or conceptualized responses within our minds. That is, on the causal view of belief acquisition we cannot make sense of how we go about revising our concepts in the light of experience of the world. Our spontaneity is reduced to a kind of partly randomized belief-reweaving machine in the manner of Dennett.

Rather, on the picture I am suggesting, what is primary is a direct awareness of the world as it appears to us, which we can then conceptualize and categorize as we see fit. We can conceptualize and categorize it in many different ways to suit our various purposes, and we can do this in several ways all at once. We can then test and modify these conceptualizations against further empirical evidence drawn from our experience of the world. This just means again, that the relation between the world and the actualization of concepts is not a passive, causal relation, but an active relation. What this also means is that in perception we bear cognitive relations to the world which are not at the same time conceptual relations.

¹² Of course, this is not all there is to the *Myth of the Given* or to foundationalist empiricism. As we have seen, the idea that our concepts are determined by their objects is a more important element of the Myth.

Now of course there are causal relations involved in perception, but these are physical processes, not logical or epistemic relations. Our epistemic relations with the world are active, involving an interplay on various levels between our perception and our conception of the world. This interplay is itself something active or conscious, to varying degrees: from the practical modification of our understanding as we find ourselves to hold a mistaken perceptual belief, to the theoretical constructs scientists create to explain empirical phenomena.

Unfortunately I do not have the space to go into more detail about this view here. What I have tried to show, however, is that given the causal view of perception, and the inferential view of justification, we cannot hope to make sense of judgement, nor of the autonomy of our conceptual schemes. To avoid both Kant's division between the phenomenal and the noumenal, and Sellars's division between the space of reasons and the space of causes, we must reconceive our notion of experience, rejecting entirely the empiricist picture to which both of these thinkers, and most since, are still partly beholden. This will enable us to relink justification and truth while maintaining the spontaneity of our conceptual scheme, thus avoiding the Hegelian synthesis towards which modern naturalist philosophy seems inexorably drawn.

REFERENCES

- MACBETH, Danielle. (2000). "Empirical knowledge: Kantian themes and sellarsian variations". *In*: *Philosophical Studies*, no. 101, p. 113-142.
- _____. "An Antinomy of Empirical Judgment: Brandom and McDowell". Forthcoming, available at: www.haverford.edu/philosophy/dmacbeth/publications/
- McDOWELL, J. (1996). *Mind and World*. Cambridge: Harvard University Press.
- RORTY, R. (1998). "The Very Idea of Human Answerability to the World: John McDowell's version of Empiricism". *In*: *Truth and Progress*. Cambridge: Cambridge University Press.
- SELLARS, W. (1997). *Empiricism and the Philosophy of Mind*. Cambridge: Harvard University Press.

CONCEIVING THE IMPOSSIBLE AND THE MIND-BODY PROBLEM

CONCEBIR LO IMPOSIBLE Y EL PROBLEMA MENTE-CUERPO

THOMAS NAGEL

University of New York, United States. thomas.nagel@nyu.edu

RECIBIDO EL 30 DE AGOSTO DE 2010 Y APROBADO EL 30 DE SEPTIEMBRE DE 2010

RESUMEN ABSTRACT

Las intuiciones basadas en la perspectiva de la primera persona fácilmente nos pueden inducir a error sobre lo que es y no es concebible. Este punto usualmente se presenta como apoyo de posiciones reduccionistas comunes sobre el problema mente-cuerpo, pero considero que se puede separar de tal perspectiva. Me parece que la fuerte apariencia de contingencia en la relación entre el funcionamiento del organismo físico y la mente consciente –una apariencia que depende directa o indirectamente de la perspectiva de la primera persona– tiene que ser una ilusión. En otras palabras, creo que hay una conexión necesaria en ambas direcciones entre lo físico y lo mental, pero que no puede descubrirse a priori. La opinión se divide fuertemente entre la credibilidad de algún tipo de reduccionismo funcionalista, y no examinaré mis razones para ubicarme en el lado antireduccionista de este debate.

Mi lectura del asunto es que nuestra falta de habilidad para obtener una concepción inteligible de la relación mente-cuerpo es un signo de la inadecuación de nuestros conceptos actuales, y que se requiere algún grado de desarrollo.

The Intuitions based on the first-person perspective can easily mislead us about what is and is not conceivable. This point is usually made in support of familiar reductionist positions on the mind-body problem, but I believe it can be detached from that approach. It seems to me that the powerful appearance of contingency in the relation between the functioning of the physical organism and the conscious mind –an appearance that depends directly or indirectly on the first person– perspective must be an illusion. In other words, I believe that there is a necessary connection in both directions between the physical and the mental, but that it cannot be discovered a priori. Opinion is strongly divided on the credibility of some kind of functionalist reductionism, and I won't go through my reasons for being on the antireductionist side of that debate.

My reading of the situation is that our inability to come up with an intelligible conception of the relation between mind and body is a sign of the inadequacy of our present concepts, and that some development is needed.

PALABRAS CLAVE KEY WORDS

a priori, primera persona, funcionalismo, problema mente-cuerpo, propiedades físicas, reduccionismo

a priori, first person, funcionalismo, mind-body problem, physical properties, reductionism

I

Intuitions based on the first-person perspective can easily mislead us about what is and is not conceivable. This point is usually made in support of familiar reductionist positions on the mind-body problem, but I believe it can be detached from that approach. It seems to me that the powerful appearance of contingency in the relation between the functioning of the physical organism and the conscious mind –an appearance that depends directly or indirectly on the first person perspective– must be an illusion. But the denial of this contingency should not take the form of a reductionist account of consciousness of the usual type, whereby the logical gap between the mental and the physical is closed by conceptual analysis –in effect, by analyzing the mental in terms of the physical (however elaborately this is done– and I count functionalism as such a theory, along with the topic-neutral causal role analyses of mental concepts from which it descends).

In other words, I believe that there is a necessary connection in both directions between the physical and the mental, but that it cannot be discovered a priori. Opinion is strongly divided on the credibility of some kind of functionalist reductionism, and I won't go through my reasons for being on the antireductionist side of that debate. Despite significant attempts by a number of philosophers to describe the functional manifestations of conscious mental states, I continue to believe that no purely functionalist characterization of a system entails –simply in virtue of our mental concepts– that the system is conscious.

So I want to propose an alternative. In our present situation, when no one has a plausible answer to the mind-body problem, all we can really do is to try to develop various alternatives one of which may prove in the long run to be an ancestor of a credible solution. This is a plea for the project of searching for a solution that takes conscious points of view as logically irreducible to, but nevertheless necessarily connected with, the physical properties of the organisms whose points of view they are. Consciousness should be recognized as a conceptually irreducible aspect of reality that is necessarily connected with other equally irreducible aspects –as electromagnetic fields are irreducible to but necessarily connected with the behavior of charged particles and gravitational fields with the behavior of masses, and vice versa. But the task of conceiving how a necessary connection might hold between the subjective and the physical cannot be accomplished by applying analogies from within

physical science. This is a new ballgame. Yet I believe it is not irrational to hope that some day, long after we are all dead, people will be able to observe the operation of the brain and say, with true understanding, "That's what the experience of tasting chocolate looks like from the outside". Of course we already know what it looks like from far enough outside: the subject taking the first reverent mouthful of a hot fudge sundae, closing his eyes in rapture, and saying "Yum".

But I have in mind some view or representation of the squishy brain itself, which in light of our understanding we will be able to *see* as tasting chocolate. While that is at the moment inconceivable, I think that it is what we would have to have to grasp what must be the truth about these matters. My reading of the situation is that our inability to come up with an intelligible conception of the relation between mind and body is a sign of the inadequacy of our present concepts, and that some development is needed. At this point, however, all one can hope to do is to state some of the conditions that more adequate concepts would have to satisfy. One can't expect actually to come up with them¹. But I shall begin by describing the present impasse.

II

When we try to reason about the possible relations between things, we have to rely on our conceptual grasp of them. The more adequate the grasp, the more reliable our reasoning will be. Sometimes a familiar concept clearly allows for the possibility that what it designates should also have features not implied by the concept itself—often features very different in kind from those directly implied by the concept. Thus ordinary prescientific concepts of kinds of substances, such as water or gold or blood, are in themselves silent with regard to the microscopic composition of those substances but nevertheless open to the scientific discovery, often by very indirect means, of such facts about their true nature. If a concept refers to something that takes up room in the spatiotemporal world, it provides a handle for all kinds of empirical discoveries about the inner constitution of that thing.

¹ My position is fairly close to that of Colin McGinn, but without his pessimism. See for example his essay "Consciousness and the Natural Order". In: *The Problem of Consciousness*. Cambridge: Blackwell, 1991. What I have to say here is also a development of a suggestion in *The View From Nowhere* (Oxford: Oxford University Press, 1986. p. 51-53).

On the other hand, sometimes a familiar concept clearly excludes the possibility that what it designates has certain features: for example we do not need a scientific investigation to be certain that the number 379 does not have parents. There are various other things that we can come to know about the number 379 only by mathematical or empirical investigation, such as what its factors are, or whether it is greater than the population of Chugwater, Wyoming, but we know that it does not have parents just by knowing that it is a number. If someone rebuked us for being closed-minded, because we can't predict in advance what future scientific research might turn up about the biological origins of numbers, he would not be offering a serious ground for doubt.

The case of mental processes and the brain is intermediate between these two. Descartes thought it was closer to the second category, and that we could tell just by thinking about it that the human mind was not an extended material thing and that no extended material thing could be a thinking subject. But this is, to put it mildly; not nearly as self-evident as that a number cannot have parents. What does seem true is that the concept of a mind, or of a mental event or process, fails to plainly leave space for the possibility that what it designates should turn out also to be a physical thing or event or process, as the result of closer scientific investigation -in the way that the concept of blood leaves space for discoveries about its composition. The trouble is that mental concepts don't obviously pick out things or processes that take up room in the spatiotemporal world to begin with. If they did, we could just get hold of some of those things and take them apart or look at them under a microscope or subject them to chemical analysis. But there is a prior problem about how those concepts might refer to anything that could be subjected to such investigation: They don't give us the comfortable initial handle on the occupants of the familiar spatiotemporal world that prescientific physical substance concepts do².

Nevertheless it is overconfident to conclude, from one's inability to imagine how mental phenomena might turn out to have physical properties, that the possibility can be ruled out in advance. We have to ask ourselves whether there is more behind the Cartesian intuition than mere lack of knowledge, resulting in lack of imagination³. Of course it

² See MCGINN, Colin. "Consciousness and Space". In: *Journal of Consciousness Studies*, 1995. no. 2, p. 220-30.

³ This is the objection that Arnauld made to Descartes, in the fourth set of objections to the *Meditations*.

is not enough just to say, “You may be mistaking your own inability to imagine something for its inconceivability”. One should be open to the possibility of withdrawing a judgment of inconceivability if offered a reason to think it might be mistaken; but there does have to *be* a reason, or at least some kind of story about how this illusion of inconceivability might have arisen.

If mental events really have physical properties, an explanation is needed of why they seem to offer so little purchase for the attribution of those properties. Still, the kind of incomprehensibility here is completely different from that of numbers having parents. Mental events, unlike numbers, can be roughly located in space and time, and are causally related to physical events, in both directions. The causal facts are strong evidence that mental events have physical properties, if only we could make sense of the idea⁴.

Consider another case where the prescientific concept did not obviously allow for the possibility of physical composition or structure –the case of sound. Before the discovery that sounds are waves in air or another medium, the ordinary concept permitted sounds to be roughly located, and to have properties like loudness, pitch, and duration. The concept of a sound was that of an objective phenomenon that could be heard by different people, or that could exist unheard. But it would have been very obscure what could be meant by ascribing to a sound a precise spatial shape and size, or an internal, perhaps microscopic, physical structure. Someone who proposed that sounds have physical parts, without offering any theory to explain this, would not have said anything understandable. One might say that in advance of the development of a physical theory of sound, the hypothesis that sounds have a physical microstructure would not have a clear meaning.

Nevertheless, at one remove, the possibility of such a development is evidently not excluded by the concept of sound. Sounds were known to have certain physical causes, to be blocked by certain kinds of obstacles, and to be perceptible by hearing. This was already a substantial amount of causal information, and it opened the way to the discovery of a physically describable phenomenon that could be identified with sound because it had just those causes and effects –particularly once further features of sound, like variations of loudness and pitch, could also be

⁴ Compare DAVIDSON, Donald. “Mental Events”. In: *Essays on Actions and Events*. Oxford: Oxford University Press, 1980.

accounted for in terms of its precise physical character. Yet it is important that *in advance*, the idea that a sound has a physical microstructure would have had no clear meaning. One would not have known how to go about imagining such a thing, any more than one could have imagined a sound having weight. It would have been easy to mistake this lack of clear allowance for the possibility in the concept for a positive exclusion of the possibility by the concept.

The analogy with the case of mental phenomena should be clear. They too occupy causal roles, and it has been one of the strongest arguments for some kind of physicalism that those roles may prove upon investigation to be occupied by organic processes. Yet the problem here is much more serious, for an obvious reason: Identifying sounds with waves in the air does not require that we ascribe phenomenological qualities and subjectivity to anything physical, because those are features of the perception of sound, not of sound itself. By contrast, the identification of mental events with physical events requires the unification of these two types of properties in a single thing, and that remains resistant to understanding. The causal argument for identification may make us believe that it is true, but it doesn't help us to understand it, and in my view, we really shouldn't believe it unless we can understand it.

The problem lies in the distinctive first-person/third-person character of mental concepts, which is the grammatical manifestation of the subjectivity of mental phenomena. Though not all conscious beings possess language, our attribution of conscious states to languageless creatures implies that those states are of the kind that in the human case we pick out only through these distinctive concepts, concepts which the subject applies in his own case without observation of his body.

They are not pure first-person concepts: To try to detach their first-person application from the third person results in philosophical illusions: e.g. that the subject of my consciousness might have been replaced five minutes ago and all my memories, personality, etc. transferred to a new subject in this same body, without any outwardly or inwardly perceptible sign –without any other physical or psychological change. If the pure first-person idea of “I” defined an individual, that would make sense, but it seems reasonably clear that the real idea of “I” has lost its moorings in this philosophical thought experiment. The point goes back to Kant, who argued that the subjective identity of the consciousness of myself at different times is not sufficient to establish the objective identity of a subject or soul⁵.

⁵ See KANT, I. Critique of Pure Reason. A 363-4: the Paralogisms of Pure Reason.

That is not to say that I understand just how the first person and the third form two logically inseparable aspects of a single concept –only that they do. This applies to all conscious mental states and events, and their properties. They are subjective, not in the sense that they are the subjects of a purely first-person vocabulary, but in the sense that they can be accurately described only by concepts in which nonobservational first-person and observational third-person attributions are logically inseparable. Such states are modifications of the point of view of an individual subject.

The problem, then, is how something that is an aspect or element of an individual's subjective point of view could also be a physiologically describable event in the brain –the kind of thing which, considered under that description, involves no point of view and no distinctively immediate first-person attribution at all. I believe that as a matter of fact you can't have one without the other, and furthermore that the powerful intuition that it is conceivable that an intact and normally functioning physical human organism could be a completely unconscious zombie is an illusion –due to the limitations of our understanding. Nevertheless those limitations are real. We do not at present possess the conceptual equipment to understand how subjective and physical features could both be essential aspects of a single entity or process.

III

Suppose, as seems likely, that I taste the smoke of my cigar when and only when my brain is in a certain physical state. What gets in the way of the thought that the experiential state of which I am introspectively aware is the physical state? The problem lies in the lack of any conceivable internal connection between a modification of my subjective point of view and a modification of the physico-chemical activity of my brain. The two may correspond extensionally as exactly as you like, but identity requires more than that. If they are the same state, it must be impossible for the one to exist without the other. And while we may have good empirical reasons to believe that that is true, the *understanding* of such an impossibility requires that the necessity of the connection between the two become intellectually transparent to us. In the case of conscious states and physiological states, it isn't just that we don't see such a necessary connection: it seems in advance that a necessary connection between two such different things is unimaginable. They seem logically unrelated.

It is very different from trying to imagine the possibility of a physico-chemical analysis of embryonic development, before one has the slightest inkling of what the analysis might be. In that case we can nevertheless understand the possibility in the abstract: there are no grounds for ruling it out a priori, since we know in general terms what it is to exhaustively analyze a largescale process into its microscopic or submicroscopic components. But that model will not help us with the mind and the brain. We are entitled to identify a familiar physical phenomenon like fire with a theoretically described process at the atomic level because the macroscopic properties of fire, including its causes and effects, can be derived from the microscopic account –can be seen to be necessitated by it. The chemical analysis of fire and air reveals, for example, that it is a necessary truth that a fire confined to a small, airtight space will quickly go out. But we cannot see how a detailed account of what is going on in the brain could exhaustively explain the taste of a cigar –not even if we could see how it explained all the physical effects of such an experience. So long as this explanatory gap remains, the identification of the states remains problematic⁶.

This does seem to call for some revision in our way of conceiving of mind, or matter, or both. The difficulty is to do this without denying what is in front of your nose. What we need is not a reductionist or eliminative revision but an expansionist one. By this I mean a conception that will permit subjective points of view to have an objective physical character *in themselves*. The reason such an expansion does not seem to me out of the question is that it doesn't involve a contradiction with the essential nature of subjective experience. Our problem is that there is no room for a necessary connection with physiology in the space of possible development defined by the concept of mind. But that does not rule out the possibility of a successor concept of mind which will both preserve the essential features of the original and be open to the discovery of such connections. This kind of thing happens all the time in the course of scientific history, as with the concepts of sound, or element, or species, or space, or number. A new concept is developed to talk about the same things as the old, one that includes most of the features of the old concept but puts them in a relation to one another and to other features that is new, and that makes it possible to see or explore further connections.

⁶ Joseph Levine has stressed this point, though I believe he is too ready to describe the problem as epistemological rather than metaphysical. See: LEVINE, Joseph. "Materialism and Qualia: The Explanatory Gap". In: Pacific Philosophical Quarterly, 1983.

Without such an expanded conception of the mental there is no prospect of overcoming the explanatory gap. But it seems to me entirely possible that a more adequate conception of subjective experience should see it as the phenomenological “inside” of certain physical processes, and that what we need is to dismantle the barriers to such a conception. If it became possible for us to think of experiences as essentially what it is like subjectively for a physical organism to be in some physiological state, the way would be open for the discovery of a posteriori necessary truths about what physiological state a particular kind of experience is. But the big step is the first one, of expanding the concept of experience by the recognition that what it contains explicitly –including its behavioral or functional implications– gives an incomplete account of the nature of experience.

The ordinary concept of water is unsaturated, so to speak, since it contains a blank space to be filled in by the discovery of the real, and essential, chemical composition of water. Just as we make room for the possibility of such discovery by denying that the manifest properties of water exhaust its nature, so we can open the possibility of an a posteriori answer to the mind-body problem by denying that the manifest properties of experience exhaust its nature. This means thinking of experiences, contrary to intuition, as events whose full nature is not revealed to experience –and more generally, thinking of the mind, contrary to Cartesian intuition, as only partially available, even in principle, to introspection. If we can do this without denying the phenomenology or reducing it to something else, we will be on the first step toward an expansionist but still nondualist response to the mind-body problem. This is so far pure fantasy, but it is the fantasy of a theoretical identification of mental events with an inner constitution that includes but is not exhausted by their introspectible or manifest character.

The search for the possible form of a theory of the relation between mind and brain has to continue, and if there can be no such theory, that too requires explanation. I believe that the explanatory gap in its present form cannot be closed –that so long as we work with our present mental and physical concepts no transparently necessary connection will ever be revealed, between physically described brain processes and sensory experience, of the logical type familiar from the explanation of other natural processes by analysis into their physico-chemical constituents. We have good grounds for believing that the mental supervenes on

the physical- i.e. that there is no mental difference without a physical difference. But pure, unexplained supervenience is not a solution but a sign that there is something fundamental we don't know. We cannot regard pure supervenience as the end of the story because that would require the physical to necessitate the mental without there being any answer to the question *how* it does so. But there *must* be a "how", and our task is to understand it. An obviously systematic connection that remains unintelligible to us calls out for a theory.

IV

The subjectivity of consciousness seems to block all reductionist proposals because, given any physicalist or functionalist description, however sophisticated, it *seems* logically possible that there should be an organism or system satisfying those conditions but nevertheless lacking any subjective point of view –a *zombie*, in current jargon. That is, there seems nothing to bar us from conceiving positively the existence of any kind of physical or functionally organized system, and then conceiving negatively that there is no subjective experience connected with it⁷. The two types of conception are so completely unrelated that the first seems incapable of ruling out the second: All we have to do is imagine the physical system from the outside, and then imagine it from the inside –as not having any inside in the experiential sense.

I said at the outset that we should be very cautious in the use of intuitions that depend on the first-person point of view, and that includes intuitions about persons or beings other than ourselves that depend on taking up their point of view in imagination. In this case the very disparity between the two forms of conception that gives rise to the strong intuition of conceivability should make us suspicious. The absence of any conceptual connection when phenomena are grasped by such disparate concepts may conceal a deeper necessary connection that is not yet conceptual because not accessible to us by means of our present forms of thought.

I still think we can rely on such thought experiments to refute the most common types of *conceptual* reductionism, however. Even if there is some kind of entailment of the mental by the physical-functional, it is not analytic or definitional: There is no hidden verbal contradiction

⁷ This argument has recently been given much prominence by David Chalmers. See: CHALMERS, David. *The Conscious Mind*. Oxford: Oxford University Press, 1996.

in the description of a zombie –even if in reality a zombie is logically impossible. There remains the question of a possible weaker conceptual link in the other direction –from the mental to the physical– through the necessity of public criteria for the application of mental concepts, which go with their distinctive first-person/third-person character. But while such public criteria may be necessary for the operation of mental concepts, they are not sufficient, and their logical relation to the reference of those concepts is not obvious. In any case, the public criteria are functional rather than physiological, on the usual interpretation, and what I am talking about here is the relation between mental states and the brain, not between mental states and behavior. Let me try to say what is wrong with conceptual arguments against the necessity of the former relation. I shall return to the private language issue later.

The following things seem *prima facie* conceivable which are pretty certainly impossible in a very strong sense, namely:

- (1) a living, behaving, physiologically and functionally perfect human organism that is nevertheless completely lacking in consciousness, i.e. a zombie;
- (2) a conscious subject with an inner life just like ours that behaves and looks just like a human being but has electronic circuitry instead of brains.

To repeat, I believe the apparent conceivability of these things reveals something about our present concepts but not about what is really possible. Analytic psychophysical reductionism is false, but there is independent reason to believe that these are not true logical possibilities, and if so, our concepts are missing something. Our current concepts don't lead to contradiction –it's not as bad as that– but they fail to reveal a logical impossibility.

Contrast these thought experiments with the *a priori* inconceivability of a number having parents. The latter involves a straightforward clash between concepts, not merely a disparity. No number could enter into the kind of biological relation with a predecessor that is a necessary condition of being a child or offspring. In that case we see a contradiction between the conditions of numberhood and the conditions of being the child of anything or anyone. In the relation of consciousness to the physical world, by contrast, our concepts fail to reveal a necessary connection, and we are tempted to conclude to the absence of any such

connection. Our intuition is of a logical compatibility, not a logical incompatibility. We conceive the body from outside and the mind from inside, and see no internal connection, only an external one of correlation or perhaps causation. But in spite of the vividness of the intuition, I believe that it reflects our conceptual limitations rather than the truth: The difference between the modes of conception is so great that there is every reason to suspect that we would be unable to see an internal necessary connection even if there were one.

Conceivability and inconceivability are the main evidence we have for possibility and necessity, but they can be misleading, and conceivability that depends on the relation between first and third person reference is particularly treacherous terrain. It may be that the physical description of the brain states associated with consciousness is an incomplete account of their essence –that it is merely the outside view of what we recognize from within as conscious experience. If anything like that is true, then our present conceptions of mind and body are radically inadequate to the reality, and do not provide us with adequate tools for a priori reasoning about them. That poses the general question of how we can attempt to develop conceptions that reflect the actual necessary connections and are therefore reliable tools for reasoning, and what determines whether there is hope of developing such concepts for a domain where we do not yet have them. After all, humans did not always have logical, geometrical, and arithmetical concepts, but had to develop them.

V

Obviously we cannot will a new conceptual framework into existence. It has to result from trying to think, in light of the evidence, about the subject we want to understand, and devising concepts that do better justice to it than the ones we have. Considered as a form of revisionism rather than analysis, the physicalist-functionalist movement in philosophy of mind might be thought to have had a similar aim, but I believe it has failed because it is too conservative: It has tried to reinterpret mental concepts so as to make them tractable parts of the framework of physical science. What is needed instead is a search for something more unfamiliar, something which starts from the conceptual unintelligibility, in its present form, of the subjective-objective link. The enterprise is one of imagining possibilities: Identity theorists like Smart, Armstrong, and Lewis tried to explain how the identity of mental with

physical states could be a contingent truth; I am interested in how some sort of mind-brain identity might be a necessary truth. That would require not only the imagination of concepts that might capture the connection, but also some account of how our existing concepts would have to be related to these and to one another. Specifically, we must think about the first-person and third-person conditions of application of mental concepts, and the physical concepts used to describe the brain. What would something be like to which *all* of these applied noncontingently?

If we think of a standard case of reference-fixing (in Kripke's sense) like "water", we observe that, prescientifically, the reference is fixed not only by contingent features of the substance, such as its being what falls from the sky as rain and fills the lakes and oceans, what comes out of the tap, and what we bathe in –but also by noncontingent manifest features, such as its density, its liquidity at room temperature and its propensity to freeze or boil under certain conditions. We can adopt a reference-fixing model (rather than the reductionist model) for the relation between functional states and mental states provided we do not think of the functional states as *contingent* reference-fixers. I think the right view is not that functional-role characterizations contingently pick out inner states whose intrinsic nature must be physically or phenomenologically specified, but rather that *all* the types of features by which we standardly identify mental states, from within and from without, are noncontingent features of those states, and that their physiological nature is *also* noncontingent.

That could be true only if they were identified with neurophysiological states whose connection, at least dispositionally, with characteristic functional roles in the organism was noncontingent. This does not seem out of the question: Integrated visual and motor systems in the brain might be part of the specification of the neurophysiological nature of color phenomenology, for example. What I am saying is not incompatible with Kripke's claim that the reference of a term like "pain" is fixed by an essential property of the referent, namely its immediate phenomenological quality –but it goes beyond this. My suggestion is that the immediate phenomenological quality, the functional role, and the physiological basis of pain are all essential properties of it, and that the apparent conceivability of their separation is an illusion.

To be precise, I conjecture that being a pain entails (nonanalytically) all three features –functional, phenomenological, and physiological– but

that only the latter two entail pain. This is because the phenomenological and the physical entail the functional (and one another) whereas the functional does not entail them. It is parallel to the case of water: There could be a watery liquid that wasn't the compound H_2O and therefore wasn't water; but being H_2O entails having the essential gross properties of water. Similarly, it is possible that there could be a state functionally equivalent to pain in a mechanism with a completely different internal constitution, but if it were both physically and phenomenologically different, it would not be the same mental state and would not be pain. Phenomenological and physiological features of mental states entail their functional features, but not vice versa.

So the proposal is that mental states may have a tripartite essence – phenomenological, functional, and physiological– but we still don't understand how this could be, since our modal intuitions go against it. In particular, we still have to deal with the apparent conceivability of an exact physical-functional replica of a conscious human being that nevertheless has no phenomenological “interior” at all –a zombie. It is clear that first-person imagination is doing the work in this thought experiment. The task of defending a necessary connection between the physical and the phenomenological would require some account of how a connection that is in fact internal remains stubbornly external from the point of view of our understanding. But I believe that can be done.

VI

Suppose I think about the taste of the cigar I am now smoking. What I must do first is to regard the experience as a state of myself of whose subjective qualities I am immediately aware, and which has certain publicly observable functional relations to stimuli and discriminatory capacities. There is already a natural illusion of contingency with respect to these latter relations, because they are concealed in my introspective identification of the experience. But this introspective identification is itself one of those mental acts that –because of the first-person/third-person link– cannot be logically separated from its functional connections (for example the capacity to distinguish this taste from that of a cigarette). Recognizing this, I can see that the Cartesian thought-experiment of imagining myself having this experience without having a body at all is an unreliable guide to what is really possible. It depends on the concealment of the necessary conditions of reference of

the phenomenological concept that I am employing to think about the experience (This is the anti-private-language point).

But now what of the relation between the experience and its physiological basis? Here I seem to be able to imagine either myself or someone else tasting exactly this flavor of cigar –and its having all the usual functional connections as well– although my brain or the other person’s brain is in a completely different state. Indeed it seems imaginable, though unlikely, that when I offer a friendly cigar to an exotic visitor from outer space who has a completely different physiology, it should taste the same to him. But here too the imagination is a poor guide to possibility, because it relies on an assumption of the completeness of the manifest conditions of reference of the concept (now taken to include functional conditions).

The first thing to acknowledge is that if there were a necessary connection between the phenomenology and the physiology of tasting a cigar, it would not be evident a priori on the basis of the ordinary concept of that experience, since the possession of that concept involves no awareness of anything about the brain. It isn’t just that, like the behavioral connections, the relation to the brain is *hidden from view* in my first-person use of the concept: The relation is completely absent from the concept, and cannot be retrieved by philosophical analysis.

Nevertheless, if there is such a relation, having the full concept (including the first person aspect) would require having a brain, indeed a brain with exactly the right physiological characteristics, and the brain would be directly involved in the act of imagination –though its involvement would be completely outside the range of my awareness in employing the concept. To imagine a mental state from the inside would be what I have called an act of *sympathetic* imagination –putting myself in a conscious state *resembling* the thing imagined– and it would be impossible to do this without putting my brain in a corresponding physical state⁸.

This shows that I cannot rely on the apparent imaginability of the separation of phenomenology and physiology to establish the contingency of the relation, since I can know in advance that this act

⁸ See NAGEL, T. “What Is It Like to Be a Bat?” *In*: *Philosophical Review*, 1974. fn. 11. This was an earlier response to the modal argument against materialism. See also HILL, Christopher: “Imaginability, Conceivability, Possibility and the Mind-Body Problem”. *In*: *Philosophical Studies*, 1995.

of imagination would seem subjectively the same whether the relation were contingent or necessary. If the relation is necessary, then I have not really succeeded in imagining the phenomenology without the physiology. The imagination here is essentially ostensive, and I cannot point to one without pointing to the other. But now, how am I to form the conception that the relation might actually be necessary –as opposed to merely acknowledging that I can't discover a priori that it isn't? I have to think that these two ways of referring –by the phenomenological concept and the physiological concept– pick out a single referent, in each case rigidly, but that the logical link cannot be discovered by inspecting the concepts directly: rather it goes only through their common link to the referent itself.

The idea would have to be, then, that there is a single event to which I can refer in two ways, both of them via concepts that apply to it noncontingently. One is the mental concept that I am able to acquire in both first and third person applications because I am a subject of this state, which has the special character of consciousness and introspective accessibility –the state of tasting a cigar. The other is a (so far unspecified) physiological concept that describes the relevant physical state of the brain. To admit the possibility of a necessary connection here, we have to recognize that the mental concept as it now operates has nothing to say about the physiological conditions for its own operation, and then consider the hypothesis of a successor concept that leaves a place for such a condition –a place that can be filled only a posteriori, by a theory of the third type of event that admits these two types of access, internal and external.

VII

What will be the point of view, so to speak, of such a theory? If we could arrive at it, it would render transparent the relation between mental and physical, not directly, but through the transparency of their common relation to something that is not merely either of them. Neither the mental nor the physical point of view will do for this purpose. The mental will not do because it simply leaves out the physiology, and has no room for it. The physical will not do because while it includes the behavioral and functional manifestations of the mental, this doesn't, in view of the falsity of conceptual reductionism, enable it to reach to the mental concepts themselves. The right point of view would be one which, contrary to present conceptual possibilities, included

both subjectivity and spatiotemporal structure from the outset, all its descriptions implying both these things at once, so that it would describe inner states and their functional relations to behavior and to one another from the phenomenological inside and the physiological outside simultaneously –not in parallel. The mental and physiological concepts and their reference to this same inner phenomenon would then be seen as secondary and each partial in its grasp of the phenomenon: Each would be seen as referring to something that extends beyond its grounds of application.

The difficulty is that such a viewpoint cannot be constructed by the mere conjunction of the mental and the physical. It has to be something genuinely new; otherwise it will not possess the necessary unity. Truly necessary connections could be revealed only by a new theoretical construction, realist in intention, contextually defined as part of a theory that explained both the familiarly observable phenomenological and the physiological characteristics of these inner events. Its character would have to be inferred from what it was introduced to explain –like the electromagnetic field, gravity, the atomic nucleus, or any other theoretical postulate. This could only be done in the context of a theory containing real laws and not just dispositional definitions, otherwise the theoretical entity would not have independent reality.

To summarize. The conjecture is essentially this: that even though no transparent *and direct* explanatory connection is possible between the physiological and the phenomenological, but only an empirically established extensional correlation, we may hope and ought to try as part of a scientific theory of mind to form a third conception that does directly entail both the mental and the physical, and through which their actual necessary connection with one another can therefore become transparent to us. Such a conception will have to be created; we won't just find it lying around. All the great reductive successes in the history of science have depended on theoretical concepts, not natural ones –concepts whose whole justification is that they permit us to replace brute correlations with reductive explanations. At present such a solution to the mindbody problem is literally unimaginable, but it may not be impossible.

REFERENCES

- CHALMERS, D. (1996). *The Conscious Mind*. Oxford: Oxford University Press, 1996.
- DAVIDSON, D. (1980). "Mental Events". *In: Essays on Actions and Events*. Oxford: Oxford University Press.
- HILL, C. (1995). "Imaginability, Conceivability, Possibility and the Mind-Body Problem". *In: Philosophical Studies*.
- LEVINE, J. (1983). "Materialism and Qualia: The Explanatory Gap". *In: Pacific Philosophical Quarterly*.
- McGINN, C. (1991). "Consciousness and the Natural Order". *In: The Problem of Consciousness*. Cambridge: Blackwell.
- _____. (1995). "Consciousness and Space". *In: Journal of Consciousness Studies*, no. 2, p. 220-30.
- NAGEL, T. (1974). What Is It Like to Be a Bat? *In: Philosophical Review*.
- _____. (1986). *The View From Nowhere*. Oxford: Oxford University Press.

MODUS VIVENDI, OVERLAPPING CONSENSUS AND STABILITY

MODUS VIVENDI, CONSENSO COINCIDENTE Y ESTABILIDAD

JHON RUDISILL

College of Wooster, United States. jRudisill@wooster.edu

RECIBIDO EL 4 DE OCTUBRE DE 2010 Y APROBADO EL 25 DE OCTUBRE DE 2010

RESUMEN ABSTRACT

En este artículo muestro como la teoría política de Hobbes, un gigante no-liberal del canon de la filosofía occidental, se puede interpretar como comprometida con alguna forma de neutralidad. Al reconocer el rol que juega la neutralidad en el pensamiento de Hobbes logramos ver que un requisito de neutralidad no es exclusivo del liberalismo. Más allá de esto, sin embargo, intento mostrar que considerar a Hobbes en este contexto revela ciertos puntos útiles de comparación con el trabajo tardío de Rawls que plantea preocupaciones sobre la viabilidad de su liberalismo político. Voy a argumentar que el liberalismo político de Rawls, mientras no ofrece una solución de *modus vivendi* a la justificación política, no está bien dotado para asegurar la estabilidad.

In this paper, I show how the political theory of a non-liberal giant of the western philosophy canon, Hobbes, can be interpreted as having a commitment to some form of neutrality. In recognizing the role neutrality plays in Hobbes's thought we come to see that a neutrality requirement is not exclusive to liberalism. Beyond this, however, I intend to show that consideration of Hobbes in this context reveals certain helpful points of comparison with Rawls's later work that raise concerns about the viability of his political liberalism. I argue that Rawls's political liberalism, while not a *modus vivendi* solution to political justification, is ill suited for the securing stability.

PALABRAS CLAVE KEY WORDS

acciones humanas, imperativo, Lógica, principios morales, reglas, pragmática, verificacionista

human actions, imperatives, Logic, moral principles, rules, pragmatics, verificationist

1. INTRODUCTION AND BRIEF BACKGROUND

“Neutrality,” many argue, is the quintessential and distinguishing trait of liberalism¹. The claim is that it is a distinctively liberal idea that moral neutrality is crucial for the legitimacy of a political order. The inference to neutrality from legitimacy comes out of a combination of what we might call *the voluntarist thesis* and *the fact of pluralism*. The voluntarist thesis holds that a political order is legitimate only if there is an acceptance of its fundamental normative principles amongst the citizens who are to live within that order. The fact of pluralism is simply the fact that the citizens in any modern state will be inescapably diverse in terms of the comprehensive normative doctrines that are endorsed. If the fundamental normative principles that serve to ground a political order are in conflict with (or, to put it differently, *fail to be neutral between*) the comprehensive normative doctrines of citizens, these citizens will not accept them and hence the order will be illegitimate. John Rawls’s later work is clearly influenced by this account of political legitimacy and the corresponding need to establish stability by securing the genuine and robust endorsement from many quite diverse persons. Rawls doubted the possibility of any such genuine and robust endorsement of a political order where the grounding normative principles are philosophically ‘thick’ and so he defended what he called “political” as opposed to “metaphysical” (“comprehensive,” “philosophical” or “substantive”) liberalism.

On Rawls’s account, political-normative principles are neutral in the appropriate and attainable senses of this concept in a way that philosophical-normative principles could never be. This is, in fact, the exact difference between these two types of principles. Political-normative principles could be located in an “overlapping consensus,” which means they are principles that would be equally acceptable to each of a group of persons who are quite diverse in terms of the

¹ Three examples of theorists who have understood neutrality and liberalism in this way are: ACKERMAN, Bruce. *Social Justice in the Liberal State*. New Haven, CO/US: Yale University Press, 1980. p. 10, DWORKIN, Ronald. *A Matter of Principle*. Cambridge: Harvard University Press, 1985. Ch. 8, and LARMORE, Charles. *Patterns of Moral Complexity*. Cambridge: Cambridge University Press, 1987. p. x. One of the most important discussions of neutrality in political philosophy is in RAWLS, John. *Political Liberalism*. Columbia: Columbia University Press, 1996. Lecture V, sections 5 & 6. The best collection of work on neutrality in political philosophy and theory is KLOSKO, George & WALL, Steven (Eds.) *Perfectionism and Neutrality: Essays in Liberal Theory*. Lanham, MA/US: Rowman and Littlefield, 2003. Recently, (May of 2008) The University of Montreal’s Centre de recherche en éthique hosted a conference entitled: “Liberal Neutrality: A Re-evaluation” featuring papers presented by Anthony Appiah, Steven Macedo, Steven Wall and Peter Demarneffe.

more comprehensive conceptions of the good and of what counts as a life worth living. In contrast, philosophical-normative principles are thought to be essentially parochial and so cannot be expected to garner widespread agreement. Rawls had hoped to support a liberal political theory without recourse to any of the typically liberal *philosophical* presuppositions² that had been called into question over and over by various critics of liberalism including Marxists, communitarians, feminists, and post-structuralists.

In this paper, I show how the political theory of a *non-liberal* giant of the western philosophy canon, Hobbes, can be interpreted as having a commitment to a form of neutrality. In recognizing the role neutrality plays in Hobbes's thought we come to see that a neutrality requirement is not exclusive to liberalism. One of the conclusions I reach, thus, is that 'neutrality' is ill suited for the purpose of identifying liberal thought and that the over reliance on this concept for this purpose has caused needless theoretical confusion. Beyond this, however, I intend to show that consideration of Hobbes in this context reveals certain helpful points of comparison with Rawls's later work that raise concerns about the viability of his political liberalism. I argue that Rawls's own version of liberalism both relies on more controversial philosophical assumptions than he acknowledged and is subject to criticisms similar to those that are raised against Hobbes.

2. HOBBS AND NEUTRALITY

In *Leviathan*, Hobbes draws upon an understanding of human nature derived primarily from observation of others and perhaps a bit of self-reflection. For Hobbes, the fact that in a state of nature we would regularly encounter non-trivial, practically problematic, and even potentially tragic conflicts, is explained by our being shortsighted egoists who lack genuine other-regarding dispositions³.

² Examples of such philosophical presuppositions include a stance on the metaphysics of the self, the nature of the good, or human nature.

³ Hobbes, in listing the passions, includes 'desire of good to another, BENEVOLENCE, GOOD WILL, CHARITY. If to man generally, 'GOOD NATURE' See: HOBBS. *Leviathan*, 6.22. This suggests that he will allow for some measure of an other-regarding disposition. Still, if there is any possibility for this sort of disposition it seems that for Hobbes it is slight. More often than not, it seems, benevolence, good will, and charity will come in the form of acts that are selfishly motivated. This is evidenced by the following of Hobbes's comments: "To give great gifts to a man, is to Honour him; because 'tis buying of Protection, and acknowledging of Power (...) To be sedulous in promoting another's good; also to flatter, is to Honour; as a signe we seek his protection or ayde" (*Ibid.*, 10.21-22).

In Chapter VI of *Leviathan*, Hobbes asserts that most of a person's appetites are "for particular things, (which) proceed from Experience, and triall of their effects upon themselves or other men"⁴. Our aversions, in contrast, are for things that we have learned by experience are hurtful and for things about which we haven't the prior experience to know if they are hurtful or not. Since experience is cumulative and changing and since no two people will have identical life experience, no person will maintain all of the same desires and aversions over her lifetime and few of us will desire all of the same things. Nevertheless, if a particular thing is desired, it is to be called "good" by the one who desires it. If a particular thing is hated, it is to be called "evil," again, by the person who hates it. These words, "good" and "evil," Hobbes tells us: "(...) are ever used with relation to the person that useth them: There being nothing simply and absolutely so; nor any common Rule of Good and Evill, to be taken from the nature of the objects themselves (...)"⁵.

Hobbes, thus, is a relativist with respect to the good and accepts a version of *the fact of pluralism*⁶. The diversity of experience and the relativistic indexing of "the good" to individuals results in there being many different (equally permissible) perspectives on what is good including what sort of life is best.

A further relevant point to note at this stage is that Hobbes is a *moral* positivist: "The Desires, and other Passions of man, are in themselves no Sin. No more are the Actions, that proceed from those Passions, till they know a Law that forbids them (...)"⁷.

Morally wrong behavior requires the institution of a sovereign legislator, which in turn, for Hobbes, requires a deliberate act of agreement on the part of those who will fall under that legislator. So, Hobbes also accepts *the voluntarist thesis*.

Due to egoism, shortsightedness and the lack of any compelling natural moral law individuals are inevitably in competition both for the instruments of power like wealth, reputation and honor but also for

⁴ *Ibid.*, 6.4.

⁵ *Ibid.*, 6.7.

⁶ I say 'a version of the fact of pluralism', because for Hobbes this pluralism is implied by the prior rejection of any naturalistic account of moral value and acceptance of moral relativism. Rawls and other political liberals would not want such thick philosophical commitments lurking in the background of their theory.

⁷ *Ibid.*, 13.10.

those particular things that are desired intrinsically. We will, Hobbes warns, “endeavor to destroy, or subdue one another”⁸. This leads us to be constantly on guard against others, afraid of their efforts to destroy us and makes it in our respective best interests to engage in preemptive strikes against one another: “(...) from this diffidence of one another, there is no way for any man to secure himself, so reasonable, as Anticipation; that is, by force, or wiles, to master the persons of all men he can, so long, till he see no other power great enough to endanger him (...)”⁹.

This leads to the insecurity and suspicion that is inevitable in the state of nature described by Hobbes as a state of war of all against all.

Out of this picture Hobbes devises a strategy for securing peace. We can design society, Hobbes maintained, to take advantage of our nature, work with that nature, to produce stability. It is rather like an engineer who figures out how to constrain a raging river, channeling its force into a highly controlled, predictable, and useful production of energy. Given our egoism and the overriding interest in self-preservation, Hobbes argues for a sovereign endowed with enormous power and efficacy. This sovereign determines and harshly enforces the laws. This provides selfish individuals (who are otherwise prone to take advantage of one another, steal from one another, frequently injure and even kill one another) with a motive to cooperate that appeals to their respective immediate, selfish desires for survival. The violation of the sovereign’s laws carries a grave risk of immanent death.

Hobbes’s thesis can be interpreted as involving a neutralist mode of justification. His system is argued to be the best, not because it will promote the interest of some one individual or type of individual, but because it alone meets the *common* and fundamental need of living free of fear of death. Hobbes recognizes that individuals will judge differently what sorts of life are desirable and no one view or group of views about the good life is the true one. He does not prioritize any of a number of competing conceptions of what counts as good because, for Hobbes, there is no controversy about these things. We all agree that the maximal promotion of selfish desires (most importantly, the desire to avoid death) is what we are each individually after. Morality, thus, is the artificial system of ‘oughts’ recommended by prudence and

⁸ *Ibid.*, 13.3.

⁹ *Ibid.*, 13.4.

enforced by the sovereign. Obedience best realizes that which we are each individually after. Instead of relying on a view about independent moral requirements governing which desires or types of desires are sinful and which aren't, Hobbes maintains only that, whatever an individual person might count as valuable, he won't see its realization if he is dead and that the only way to sufficiently minimize the risk of premature death is to contract into a state that will be governed by an all-powerful and indivisible sovereign. The fundamental normative principles grounding Hobbes's preferred political order, therefore, are neutral between the indefinite number of actual and possible conceptions of what, more specifically and idiosyncratically, is and can be counted as desirable.

3. LEVIATHAN AS A *MODUS VIVENDI* SOLUTION AND THE ARGUMENT AGAINST *MODUS VIVENDI* SOLUTIONS

The sort of neutrality that I see in Hobbes's *Leviathan* is that of a mere *modus vivendi*. Rawls gives the following illustration of the idea of a mere *modus vivendi*:

A typical use of the phrase "modus vivendi" is to characterize a treaty between two states whose national aims and interests put them at odds. In negotiating a treaty each state would be wise and prudent to make sure that the agreement proposed represents an equilibrium point: that is, that the terms and conditions of the treaty are drawn up in such a way that it is public knowledge that it is not advantageous for either state to violate it. The treaty will then be adhered to because so doing is regarded by each as in its national interest including in its interest in its reputation as a state that honors treaties¹⁰.

The difficulty with a *modus vivendi* approach is its fragile stability. In such an approach, points out Rawls, "social unity is only apparent, as its stability is contingent upon circumstances remaining such as not to upset the fortunate convergence of interests"¹¹. It makes no sense, according to Hobbes, to speak of *moral oughts* in the state of nature. But, there are facts in the state of nature about what will increase and decrease the probability of desire satisfaction and so there are *prudential oughts*.

¹⁰ RAWLS, J. *Political Liberalism*. New York: Columbia University Press, 1996. p. 147.

¹¹ *Ibidem*.

urthermore, for Hobbes, “rationality” is just recognizing and pursuing whatever is in one’s interest where ‘interest’ is understood in terms of egoistic desire-satisfaction. Consequence of this is that, for Hobbes (unlike, for example, Plato) the question “why ought I do what I morally ought to do?” is not trivial. It is possible that prudential prescriptions and moral prescriptions could diverge. Their alignment is a matter of contingent circumstance and not conceptual necessity.

When circumstances change, there is nothing about a *modus vivendi* to recommend continued adherence. Hobbes, is aware of this but thinks that he is able to avoid the difficulty in part because of (a) the consistency of human nature and (b) the awesome power of his sovereign. That he is, as I am convinced, wrong in his view of human nature virtually guarantees that the Leviathan variety *modus vivendi* is unstable. Many will find life in such state worse than death and so will revolt. Beyond this, it is reasonable to suppose that no sovereign can be powerful and informed enough that a subject could *never* violate its laws so as to gain tremendously without risk of being caught. If a subject were ever in such circumstances there would be nothing about the *modus vivendi* that could obligate or motivate him to refrain from violating the law.

4. THE SUPPOSED STABILITY OF AN OVERLAPPING CONSENSUS

Rawls’s argument that *political liberalism*, does not amount to a *modus vivendi* is as follows. The state and its institutions are justified only with reference to certain basic ideas that are located in the consensus that is the overlap of multiple comprehensive conceptions of the good. There is no claim as to the *truth* of any of these basic ideas or to the truth of any substantive philosophical commitments that entail them. There is simply a fortunate, historically produced, convergence of comprehensive conceptions. Certainly, any one of these conceptions is loaded with philosophically substantive and so controversial commitments. For each person (each “endorser of a comprehensive conception”) you can say that *her* commitment to the ideas in the overlap derives from the comprehensive conception *she* endorses. So each of us has his or her own reasons (drawing upon his or her own comprehensive doctrine) for supporting the state that is justified with reference to and designed in a way governed by, the basic ideas in the overlap. But the justificatory force of these ideas rests upon their being shared, not

upon any presupposition about the truth or superiority of any one of the comprehensive doctrines. *I* will support the state that is justified in this way because *I* accept these justifying principles for *my own* reasons. *You* support the state because *you* accept these justifying principles for *your own* reasons. "An overlapping consensus, therefore, is not merely a consensus on accepting certain authorities, or on complying with certain institutional arrangements, founded on a convergence of self- or group interests. All those who affirm the political conception start from within their own comprehensive view and draw on the religious, philosophical, and moral grounds it provides"¹².

Since we each have our own reasons, stemming from our own comprehensive doctrine, we will remain committed to the political conception even if circumstances change such that our own comprehensive doctrine becomes dominant. Thus, Rawls believes, we have stability.

5. CONCLUSION: WHY THE OVERLAPPING CONSENSUS IS NOT STABLE

Here is the problem. Suppose we find our diverse comprehensive doctrines 'overlapping' in the way Rawls suggests. This fact is historically contingent. So long as the diverse citizens continue to endorse *these doctrines*, and avoid significant alterations to them by mixing in dogmatism or sectarianism (rendering them unreasonable) we could expect stability. But this stability is as tenuous as that provided by Hobbes' *modus vivendi*¹³. Somewhat ironically, political liberalism can hasten the spread of dogmatism and ultimately set the stage for a loss of stability. The method of avoidance of substantive philosophical matters that is at the core of *political liberalism* is also the internal seed of its own undoing.

A state that is given such an austere and minimalist justification as that which can be squeezed out of an overlapping consensus runs a serious risk of degenerating into an anemic and listless, non-progressive

¹² *Ibid.*, p. 148.

¹³ For a distinct criticism of Rawls as facing the dilemma of either having to settle for a *modus vivendi* himself or commit to some form of comprehensive liberalism, see TALISSE, Robert B. *Democracy After Liberalism: Pragmatism and Deliberative Politics*. London: Routledge, 2005. p. 55-63.

and stagnant collectivity. Because the political liberal's state eschews any substantive commitments to the *truth* of any claims about human nature and the good for humans and relies instead only on there being a few commonly held ideas (none of which needing to be either true or dependent upon more fundamental notions that need to be true) it's own justificatory structure can hardly avoid dogmatism.

Rawls attempts to identify a small number of basic ideas that are common to several diverse comprehensive conceptions. He then uses them to justify a particular sort of liberal state. However, the justificatory force of these ideas is independent of their truth, resting only on the fact that they are shared. There is no need to question these ideas or to consider why they are preferable to some other ideas located only in those excluded conceptions of the good that do not overlap. The particular "political values" that Rawls relies on to justify the liberal state will, after a time, slip into the status of what John Stuart Mill refers to as "dead dogma". For the purposes of the state –the shape of its institutions and character of its policies– no independent reasons or philosophical grounds are required in support of these political values. More than this, however, independent reasons and philosophical grounds are *positively excluded* from consideration since it is inevitable that there will be disagreements between individuals in terms of what their independent reasons are in support of these political values. These disagreements are inevitable because the support of the overlapping political values is provided for each person by his or her comprehensive conception.

The political liberal's likely response to this concern is to say that there is nothing in his theory that prevents *individuals* from this sort of substantive reflection regarding, and critical scrutiny of, their respective comprehensive conceptions and the ideas found in their overlap. The difficulty with this response, though, is that it indicates the extreme degree to which political liberalism privatizes the issues most important and fundamental to the determination of a just state. Consequently, the quality of scrutiny regarding such issues will be low. The mere fact that political liberalism does not explicitly *prevent* such reflection and scrutiny does not mean that there will likely be such reflection and scrutiny and there is a reasonable concern that there likely won't be. The quality of whatever scrutiny exists is bound to be low because it will be an activity relegated to the private sphere and so no criteria of public justification need be met. The expected result will be an ever-increasing frequency of individuals endorsing doctrines that are sectarian and that, thus, do

not overlap. Once the overlap disappears, or even becomes an overlap of conceptions endorsed by a small number, stability will be lost.

REFERENCES

ACKERMAN, Bruce. (1980). *Social Justice in the Liberal State*. New Haven, CO/US: Yale University Press.

DWORKIN, Ronald. (1985). *A Matter of Principle*. Cambridge: Harvard University Press.

LARMORE, Charles. (1987). *Patterns of Moral Complexity*. Cambridge: Cambridge University Press.

RAWLS, J. (1996). *Political Liberalism*. New York: Columbia University Press.

KLOSKO, G. & WALL, S. (Eds.). (2003). *Perfectionism and Neutrality: Essays in Liberal Theory*. Lanham, MA/US: Rowman and Littlefield.

TALISSE, Robert. B. (2005). *Democracy After Liberalism: Pragmatism and Deliberative Politics*. London: Routledge.

FREGE AND NUMBERS AS SELF-SUBSISTENT OBJECTS*

FREGE Y LOS NÚMEROS COMO OBJETOS AUTO-SUBSISTENTES

GREGORY LAVERS

Concordia University, Canada. glavers@alcor.concordia.ca

RECIBIDO EL 2 DE OCTUBRE DE 2010 Y APROBADO EL 25 DE OCTUBRE DE 2010

RESUMEN

En este artículo se argumenta que Frege no es el metafísico platónico sobre matemáticas que se considera normalmente. Se muestra que el proyecto fregeano tiene dos etapas distintas: la identificación de lo que es verdadero en nuestras nociones ordinarias, y luego la provisión de una explicación sistemática que comparte los aspectos identificados. Ninguna de estas etapas involucra mucha metafísica. El artículo critica en detalle la interpretación que hace Dummett de los párrafos §§55-61 del *Grundlagen*. Estas secciones están bajo el encabezado 'Todo número es un objeto auto-subsistente' y Dummett las describe como las que contienen los peores argumentos planteados por Frege. Se arguye que, esencialmente, todos los puntos interpretativos de Dummett son erróneos. Finalmente, muestro que los planteamientos de Frege sobre la independencia de las matemáticas con respecto a los humanos y sus actividades tampoco lo comprometen con ninguna posición metafísica particular.

PALABRAS CLAVE

Dummett, Frege, *Grundlagen*, independencia de las Matemáticas, platonismo metafísico.

ABSTRACT

This paper argues that Frege is not the metaphysical platonist about mathematics that he is standardly taken to be. It is shown that Frege's project has two distinct stages: the identification of what is true of our ordinary notions, and then the provision of a systematic account that shares the identified features. Neither of these stages involves much metaphysics. The paper criticizes in detail Dummett's interpretation of §§55-61 of *Grundlagen*. These sections fall under the heading 'Every number is a self-subsistent object' and are described by Dummett as containing the worst arguments put forward by Frege. It is argued that essentially all of Dummett's interpretive points are mistaken. Finally, I show that Frege's claims about the independence of mathematics from humans and their activities does not commit him to any particularly metaphysical position either.

KEY WORDS

Dummett, Frege, *Grundlagen*, independence of Mathematics, metaphysical platonism.

* © 2010 Gregory Lavers.

INTRODUCTION

Frege is often described as perhaps the most prototypical realist in the philosophy of mathematics. It is important, however, to clearly separate what Frege says in terms of identifying desiderata for a systematic treatment of number from his remarks about how such a systematic treatment is to proceed. I contend that when this is done, what emerges is a picture of Frege as someone quite unconcerned with *metaphysical* issues. Neither of the two stages in providing a systematic account of arithmetic is particularly metaphysical. I hope to show that Frege is largely unconcerned with establishing metaphysical theses of any kind.

The present paper has three sections. In the first I give a definition of metaphysical realism that ties the question of whether one is engaged in metaphysical issues to the aims of the project. I look at Frege's statement of his aims and show that they do not support a metaphysical reading of his program. In the second section, I analyze in detail Dummett's remarks (Dummett, 1991) on Frege's arguments in §§55-61 of *Grundlagen*. I argue that Dummett's interpretation is implausible, and implausible precisely because he interprets Frege as a metaphysical realist. In the third section, I turn to the question of Frege's remarks about the independent existence of mathematical objects and the independent truth of mathematical propositions. I argue that, here as well, there is nothing about these remarks that is particularly *metaphysical*.

1. FREGE'S AIMS AND METAPHYSICAL REALISM

To properly understand Frege's attitude towards the question of realism we need to look first at his goals. By carefully examining Frege's stated aims, it becomes apparent that he is not the metaphysical realist that he is standardly taken to be. In fact, I believe that the connection between one's overall goals and one's stance on realism is so strong that one can be defined in terms of the other. That is, I will take it as definitional of the position of *metaphysical realism* that anyone who holds this position has as their overarching goal to describe how reality is in some ultimate sense¹. In defining *metaphysical realism* in this way, I do not take myself

¹ Erich Reck, in his *Frege's influence on Wittgenstein: reversing metaphysics versus the context principle*. (In: TAIT, W. W. (Ed.). *Early Analytic Philosophy: Frege, Russell, Wittgenstein*. (pp. 123-85). Chicago and La Salle, IL: Open Court, 1997), argues that Frege is not a metaphysical realist in an argument that likens Frege's use of the context principle to the philosophy of the later Wittgenstein.

to be diverging from its standard sense. One is a metaphysical realist if, with your account –or of number, say– you try to correctly picture how the world is in some ultimate sense. To take a particularly fitting example, if one is giving an account that holds that numbers are objects, and in fact extensions of a certain kind, then that account is acceptable only if numbers *really are* particular kinds of extensions². So, if being a metaphysical realist is a question of what one's ultimate goals are, then examining what one says about one's goals can help determine if one is a realist. Let us now, then, examine Frege's remarks concerning the goals of his project:

In the introduction to *Grundlagen*, Frege describes his project by saying: "I realize that, as a result, I have been led to pursue arguments more philosophical than many mathematicians may approve; but any thorough investigation of the concept of number is bound always to turn out rather philosophical. It is a task that is common to mathematics and philosophy"³.

We see here Frege saying that he was forced by the nature of the subject matter to be somewhat philosophical. So, this already calls into question any view that holds that *from the outset* Frege sought to defend a position on the *metaphysical* status of arithmetic. He continues describing his project by saying: "Even I agree that definitions prove their worth by being fruitful"⁴. Here 'Even I' is in reference to Frege's standards of proof, and is not a comment on his realism. Again: "Definitions show their worth by proving fruitful. (...) Let us try, therefore, [to see] whether we can derive from our definition of the Number which belongs to the concept F any of the well known properties of number"⁵ or again consider: "Now our concern here is to arrive at a *concept of number usable for the purpose of science*; we should not therefore, be deterred by the fact that in the language of everyday life number appears also in attributive constructions. That can always be got round"⁶.

² Being a metaphysical realist in this sense is closely related to holding a correspondence theory of truth. See RICKETTS, T. "Objectivity and objecthood: Frege's metaphysics of judgement". In: HAAPARANTA, L. & HINTIKKA, J. (Eds.). *Frege Synthesized*. (pp. 65-95). Dordrecht: Reidel, 1986; "Frege on logic and truth". In: *Proceedings of the Aristotelian Society Supplement*, 1996. vol. 70, p. 121-140 and RECK, E. H. "Frege on truth, judgement, and objectivity". In: *Grazer Philosophische Studien*, 2007. vol. 75, no. 1, p. 149-173, for convincing arguments that Frege did not see truth in this way.

³ FREGE, G. *The Foundations of Arithmetic*. 2d. ed. Evanston, IL: Northwestern University Press, 1980. p. v.

⁴ *Ibid.*, p. ix.

⁵ *Ibid.*, §70.

⁶ *Ibid.*, §57, my emphasis.

What, then, can we say about Frege's goals? Frege believes that the goal in providing an analysis of number is to put forward a definition that is "usable for the purpose of science". This condition is met if, from the definition, the ordinary properties of number are recoverable. So here we have what we can call Frege's first desideratum concerning a definition of number. Let us call *this the desideratum of standardness*.

A second desideratum becomes apparent when Frege provides his own definition of numbers as extensions of second-order concepts. He states "That this definition is correct will perhaps be hardly evident at first. For do we not think of the extensions of concepts as something quite different from numbers? (...) [I]t is not usual to speak of a Number as wider or less wide than the extension of a concept; but neither is there anything to prevent us from speaking in this way, if such a case should ever occur"⁷.

Here we see Frege expressing some tolerance toward a definition of number. The definition need not match up exactly with our ordinary notion. The definition might confer on numbers properties that they did not previously have. Frege believes that the new properties conferred on number by his definition are harmless in the sense that they will arise so seldom as to not significantly affect the practice of arithmetic. This then can be seen as Frege's second desideratum for an account of arithmetic. That is, an account should not assign new properties to number in such a way that it will have a serious impact on the practice of mathematics. Let us call this desideratum *the desideratum of harmlessness*.

We have now identified Frege's two desiderata for a definition of number. A definition should allow us to recover the properties we usually associate with number. It should also not introduce new properties that would alter the practice of mathematics. When he is criticizing competing definitions, it is clear that he is applying the same standards. Consider one of his earliest criticisms of a rival view that is presented in the *Grundlagen*: "When Stricker, for instance, calls our ideas of number motor phenomena and makes them dependent on muscular sensations, no mathematician can recognize his numbers in such stuff or knows what to make of such propositions"⁸.

⁷ *Ibid.*, §69.

⁸ *Ibid.*, p. v.

The problem with Stricker's definition is that number as he defines it has essentially nothing in common with number as we ordinarily conceive it. Again consider Frege on a psychologistic accounts of number:

If the number two were an idea, then it would have straight away to be private to me only. Another man's idea is, *ex vi termini*, another idea. We should then have it might be many millions of twos on our hands. We should have to speak of my two and your two, of one two and of all twos. If we accept unconscious ideas, we should have unconscious twos among them, which would return subsequently to consciousness. As new generations of children grew up, new generations of twos would continually be born, and in the course of millennia these might evolve, for all we could tell, to such a pitch that two of them should make five. Yet in spite of all of this, it would still be doubtful whether there existed infinitely many numbers, *as we ordinarily suppose*. 10^{10} , perhaps, might be only an empty symbol, and there might exist no idea at all, in any being whatever, to answer to that name⁹.

One might interpret this as a straightforward *reduction ad absurdum*. On such an interpretation Frege's goal here is to show that *it is false* that numbers are ideas by showing that it leads to absurdity. However, there is a problem with this interpretation. If this were his goal here, then when he presents his own definition of numbers as extensions, he would need to provide an argument to the effect that *this is true*. Of course, he provides no such argument. In fact, he says that although it might be somewhat artificial to take numbers to be extensions, there is nothing wrong with this. In light of the two identified desiderata, another interpretation becomes evident. Frege is not concerned with the question of whether numbers *really are ideas or not*. Frege is rejecting attempts to define numbers psychologically precisely because such a definition would satisfy neither the desideratum of standardness, nor the desideratum of harmlessness. In fact the failure of one or the other of these desiderata is argued for in each sentence in the above quote.

Does what I have said so far show that Frege is not a realist? That is not what I want to claim at all. Frege takes numbers to be objects and takes mathematical claims to express objective truths. So, he is clearly a realist in a certain sense. What I am claiming is that there is nothing

⁹ *Ibid.*, §27, my emphasis.

metaphysical about his realism. When Frege claims that numbers are objects and mathematical claims express objective truths, he is not making a metaphysical pronouncement. What he is doing is stating which features an account of number must have in order to satisfy what I have called his desideratum of standardness. Frege takes these to be properties of numbers as we ordinarily conceive of them, and therefore properties that ought to be preserved by an account of number. He does not take this to make any metaphysical claim about the ultimate nature of reality. In fact he warns against reading his claim that numbers are self-subsistent objects in this way: "The self-subsistence which I am claiming for number is not to be taken to mean that a number word signifies something when removed from the context of a proposition, but only to preclude the use of such words as predicates or attributes, which appreciably alters their meaning"¹⁰.

In this section I have tied the question of the degree to which an account of arithmetic is realist in a metaphysical sense to the expressed goals of the account. If the goal of the account is to reflect the ultimate nature of reality (or at least those features that concern the subject matter of the account), then it is a metaphysically realist position. We have seen that Frege's stated goals are actually quite pragmatic. Frege wants to give an account of arithmetic "suitable for the purposes of science". Although he claims, for instance, that numbers are self-subsistent objects, he does not mean for this to correctly picture how things in reality actually are in any metaphysical sense. As Ricketts (1986), Ricketts (1996), and Reck (2007), for instance, argue, Frege does not accept anything like the correspondence theory of truth.

Even if, however, Frege's ultimate goal was not to give a theory that corresponds to the way the world is in itself, the question of truth cannot be completely avoided. Truth is a concept of central importance for Frege. In a language like that of *Begriffsschrift*, all of the provable judgements express truths. How is the quite pragmatic reading of Frege defended here consistent with the claim that provable judgements in Frege's systematic languages express truths? To answer this we need to explore Frege's discussion of the paradox of analysis. The paradox of analysis states that no analysis can be both informative and correct. If we say what it is to be *A* is to be *B*, then there are two possibilities. If '*B*' has exactly the same meaning as '*A*', then the analysis is uninformative.

¹⁰ *Ibid.*, §60.

If 'B' means something different, then the account is incorrect. Frege provided several attempts at a solution to the paradox of analysis, but I will, here, focus on his final solution to the paradox¹¹. Frege thought that even the pre-systematic use of scientific terms, presumably including 'number', lack a completely clear meaning: "Now it may happen that this sign (word) is not altogether new, but has already been used in ordinary discourse or in a scientific treatment that precedes the truly systematic one. As a rule, this usage is too vacillating for pure science"¹².

If even the scientific treatment of concepts such as 'number' has a use that is too vacillating, how is an analysis supposed to work? The answer is provided in 'Logic in Mathematics':

If we have managed in this way to construct a system of mathematics without the need for the sign A [a sign with an existing sense], we can leave the matter there; there is no need at all to answer the question concerning the sense in which –whatever it may be– this sign had been used earlier. In this way we court no objections. However it may be felt expedient to use the sign A instead of the sign B [the newly introduced sign]. But if we do this, we must treat it as an entirely new sign which has no sense prior to the definition. We must explain that the sense in which this sign was used before the new system was constructed is no longer of any concern to us, that its sense is to be understood purely from the constructive definitions that we have given. In constructing the new system we take no account, logically speaking, of anything in mathematics that existed prior to the new system. Everything has to be made anew from the ground up. Even anything we accomplish by our analytical activities is to be regarded only as preparatory work which does not itself make any appearance in the system itself¹³.

¹¹ Frege ("Review: Husserl, philosophy of arithmetic". In: MCGUINNESS, B. (Ed.). *Collected Papers on Mathematics, Logic and Philosophy* (pp. 293–340). Oxford: Basil Blackwell, 1984a) discusses the paradox explicitly (although, of course, not by that name). Michael Beaney discusses Frege's views on analysis in BEANEY, M. *Frege: Making Sense*. London: Duckworth, 1996. chap 5 and his "Carnap's conception of explication: from Frege to Husserl?" (In: AWODEY, S. & KLEIN, C. (Eds.). *Carnap Brought Home: The View from Jena* (pp.117–150). Chicago and La Salle, IL: Open Court, 2004) including looking at his earlier attempts to avoid the paradox. Joan Weiner discusses problems related to Fregean analyses in WEINER, J. "What's in a numeral? Frege's answer". In: *Mind*, 2007. vol. 116, no. 463, p. 677–71).

¹² FREGE, G. "On the foundations of geometry: second series". In: MCGUINNESS, B. (Ed.) *Collected Papers on Mathematics, Logic and Philosophy* (pp. 293–340). Oxford: Basil Blackwell, 1984b. p. 302–302.

¹³ FREGE, G. "Logic in mathematics". In: HANS HERMES, F. K. & KAMBARTEL, F. (Ed.). *Posthumous Writings* (pp. 203–250). Oxford: Basil Blackwell, 1979. p. 211.

We see here Frege's final view on how an analysis is supposed to function. We begin by identifying what is true of the notion prior to its truly systematic treatment. This, however, is only the first step, and is a somewhat vague matter. We then replace the old term with a new one (or retain the old one *as the new one* for convenience). The new term's meaning, however, is to follow exclusively from the definition –ties to the old notion are severed. As I have explained Frege's aim in *Grundlagen*, he is engaged very much in a project along these lines. He does not argue that numbers really are and always have been extensions. He offers his definition as a replacement for our ordinary notion of number.

At this point, it will evidently be charged that I am reading Frege's remarks from 'Logic in mathematics' back into *Grundlagen*, despite the two works being separated by thirty years. Of course, Frege did not have his views on analyses expressed in 'Logic in mathematics' worked out at the time of *Grundlagen*. This is evidenced by his formulating several intermediate positions on analysis in the intervening years. However, even if Frege did not have *his general views* on explication worked out by the time of *Grundlagen*, what he does in *Grundlagen* is nonetheless *an instance of* such an analysis. Frege makes no attempt to argue that numbers always were extension. He thinks that number as he defines it is sufficiently similar to the ordinary notion that it could be used in its place. In this section I have argued for an interpretation of Frege where his aims are more practical and less philosophical than standardly assumed. In the next section I turn to an interpretation at odds with this one.

2. DUMMETT ON §§55-61

Michael Dummett has defended a strongly realist interpretation of Frege. In fact, Dummett's interpretation amounts to an interpretation of Frege as metaphysical realist in exactly the sense defined in the last section. Dummett takes Frege to hold that an account of arithmetic ought to reflect how things are in the world –and in particular the ontology of the world. Dummett writes "Frege's realism was not the most important ingredient in his philosophy: but the attempt to interpret him otherwise than as a realist leads only to misunderstanding and confusion"¹⁴ I do not want to deny that Frege was a realist. Frege believed that numbers

¹⁴ DUMMETT, M. "Frege as a realist". In: SLUGA, H. (Ed.). *Meaning and Ontology in Frege's Philosophy*, Vol. 3 of *The Philosophy of Frege* (pp. 109-122). New York & London: Garland, 1993. p. 109.

are objects, and that numerical statements express objective truths. But unlike Dummett I don't see these as Frege's conclusions. Rather, these express desiderata concerning an account of number. So Frege certainly was a realist – just not a metaphysical realist. The measure of an account is not whether it correctly reflects the true structure of metaphysical reality, but that it preserve the properties that we standardly take numbers to have. In this section I wish to show that my interpretation of Frege's realism allows for a clearly more appropriate interpretation of §§55-61. Dummett (1991) describes §§55-61 of *Foundations* (the sections that fall under the heading 'Numbers are self-subsistent objects') as containing some of the worst philosophical arguments presented by Frege¹⁵.

Dummett discusses these sections at length in the chapter of *Frege: Philosophy of Mathematics* titled 'Two strategies of analysis'. Here, Dummett distinguishes between two strategies for defining number. On the one hand, we could be radical substantialists and begin by defining numerical terms and then translate adjectival uses of number into claims involving numerical terms. On the other hand, we can define adjectival uses of number and then translate all statements which involve numerical terms into ones that do not. This second strategy Dummett calls the *radical adjectival strategy*. In §55, as Dummett understands him, Frege proposes three definitions that pursue a radical adjectival strategy. That is, the definitions attempt to define our use of numbers as adjectives without the use of numerical terms. The definitions are stated as follows:

the number 0 belongs to a concept, if the proposition that a does not fall under that concept is true universally, whatever a may be.

the number 1 belongs to a concept F, if the proposition that a does not fall under F is not true universally, whatever a may be, and if from the propositions that "a falls under F" and "b falls under F" it follows universally that a and b are the same.

¹⁵ Wray, K. B. focusses exclusively on Dummett's interpretation of x56. Although I agree with Wray that Dummett seriously misreads Frege here, I don't see that Wray has identified what exactly it is about Dummett's interpretation that is at fault. I do not think it is simply that "Frege is not trying to show that the adjectival strategy cannot make sense of our use of number statements, as Dummett suggests, but rather, his intention is to show that only the substantialist strategy can provide us with an adequate definition of number" ("Reinterpreting section 56 of Frege's *The Foundations of Arithmetic*". In: *Auslegung: A Journal of Philosophy*, 1995. vol. 20, no. 2, p. 79). The main reason I don't think Wray's interpretation is acceptable is that it is not altogether clear how it differs from Dummett's. Presumably, if only a substantialist view will be acceptable, then this itself is a strong argument against the adjectivalist.

the number $(n+1)$ belongs to the concept F , if there is an object a falling under F and such that the number n belongs to the concept "Falling under F , but not a "

However, on Dummett's view:

In stating them, [Frege] makes heavy use of his jargon. Instead of saying 'There is just 1 F ', (...) he says 'The number 1 belongs to the concept F '. This, of course, obscures the fact that these are *adjectival* uses of number words that he is defining¹⁶.

In §56 Dummett sees Frege as objecting to the adjectival definition of number just presented in §55. Frege presents two arguments here. The first involves the use of the variable n in the third definition. If numerical terms have not been defined, but only phrases containing numerical terms, then we are not free to use a variable that ranges over numbers. Dummett recognizes this as a valid criticism of Frege's proposed definition. Dummett argues, however, that if Frege had taken the adjectival strategy more seriously he would have realized that this objection is not valid against all attempts to give an adjectival definition of number. As a result, Dummett spends much of the chapter providing an adjectival definition that avoids this difficulty while at the same time involving no numerical terms. Frege mentions, as an illustration of this problem with the third definition, that the proposed definition does not rule out Julius Caesar as a number. Dummett can make no sense of this; he merely states that no one reading the work for the first time could see this as relevant, and then he moves on.

The second objection that Frege puts forward is that the definitions do not allow us to show that if a and b are both numbers belonging to a concept F , then $a = b$. Concerning this, of course, Dummett points out that identity, for Frege, is a relation between objects. As such, Dummett sees this argument as completely question begging. As with his handling of the first objection, Dummett shows that with proper definitions this problem can be avoided. He does this by defining a relation between numerically definite quantifiers that can be used to show that only one numerically definite quantifier applies to a given concept. Concerning §56, Dummett writes: "Frege aimed to show all

¹⁶ DUMMETT, M. Frege: Philosophy of Mathematics. Cambridge, MA: Harvard University Press, 1991. p. 100.

three definitions erroneous, and thereby to prove a purely adjectival strategy unfeasible, because numbers have to be recognized as being objects. In this he utterly failed: in fact §56 may be stigmatized as the weakest in the whole of *Grundlagen*. The arguments lack all cogency: they more resemble sleight of hand"¹⁷.

Dummett sees §57 as merely a continuation of §56. Therefore, he sees it as just as question begging as the previous section. In addition to this, Dummett finds Frege to be self-undermining. Here Frege shows how we can translate adjectival uses of number into statements involving singular terms standing for numbers. But if we are free to reinterpret in one direction, why should we take the surface syntax seriously in the other? Dummett concludes that Frege has given no good argument that numbers are objects. Concerning §§58-61, Dummett has no objections. He merely points out that these sections contain quite admirable defense of the claim that numbers are objects against the possible objections that we can have no ideas of them or that they have no spatial location.

I seriously disagree with almost every aspect of Dummett's interpretation of §§55-57. In the previous section, I argued that Frege is not a metaphysical realist. The way that Dummett interprets §§55-57 is possible only on the assumption that Frege is a metaphysical realist. Dummett thinks that Frege is concerned to show that numbers really are objects, and sees §§55-57 as a sustained, but incredibly poor, argument to this effect. Dummett is here reading far too much into the text.

Dummett interprets Frege as proposing for consideration radical adjectival definitions of number in §55, and in so doing, using incredibly questionbegging jargon. However, there is nothing about §55 which suggests that Frege is considering a radical adjectival definition of number. After presenting his 'fundamental thought' in §46 and discussing it in the following sections, Frege says that this leads almost immediately to the definitions of §55¹⁸. In all three of these definitions, Frege includes what are clearly meant to be terms that stand for numbers (as objects). If one interprets these definitions as following the radical adjectival strategy, as Dummett does, the use of the definite articles that

¹⁷ *Ibid.*, p. 105.

¹⁸ On page 5 of Frege *The Basic Laws of Arithmetic: Exposition of the System* (Edited and translated by Montgomery Furth. Berkeley and Los Angeles: University of California Press, 1967), Frege describes the claim that statements of numbers contain an assertion about a concept as the fundamental result of *Grundlagen*.

appears in each of the three definitions must be chalked up to either *a major oversight or major dishonesty*. It is not clear which option is meant to be implied by Dummett's phrase 'question begging jargon'. If on the other hand, we assume that Frege chose consciously and without dishonesty to use the definite article in the definitions, then §55 contains a discussion of readily suggested but ultimately unacceptable definitions of *numerical terms*. Frege understands 'the number n belongs to F ' as having the form $B(n, F)$, where B (the 'belongs to' relation) is a relation between an object and a first level property. On this interpretation, §55 does not contain radical adjectival definitions *at all*. Instead, what Frege is considering here is introducing numerical terms via the role they play in numerically definite quantifiers.

What then of §56? Of course, since §55 does not contain a radical adjectival definition of number, §56, which discusses these definitions, does not consider a radical adjectival strategy either. What Frege considers in this section is whether the numerical terms have been sufficiently well-defined. He concludes that they have not. The first reason given involves the introduction of the variable n ranging over the numbers in the third definition. His objection to this is that the variable ranging over numbers is illegitimate since the numbers have not yet actually been defined. To support this last point Frege mentions the Julius Caesar problem. Dummett, as we saw, says little of the mention of Caesar here other than that it is baffling. This is not surprising. If the definitions of §55 are not meant to contain *numerical terms*, then the substitution of Julius Caesar for a number *does not make any sense at all*. But on an interpretation where the definitions in §55 are meant to define numerical terms –that is terms for numbers as objects– it makes perfect sense. If numbers are objects, then a definition of number should tell you which objects are numbers and which are not. Dummett may be right that the mention of Caesar may be completely unexpected by first time readers, but as surprising as it maybe, on the interpretation defended here it is perfectly well motivated.

Frege then points out that it also does not follow from the definitions that if a and b are both numbers that belong to the same concept, that $a = b$. Dummett, in holding that the radical adjectival strategy is under consideration, sees Frege, by considering identities, as once again 'taken in by his own jargon' and assuming what he is trying to prove. However, if the definitions in §55 were never meant to be radical adjectival definitions, then a far more charitable interpretation is possible. If Frege

intended the definitions in §55 as definitions of numerical terms, then Frege is merely pointing out that the definitions do not allow us to derive a basic property of numbers –that for any (sortal) concept there is one and only one number of things that falls under it. Remember, we saw in the previous section that deriving the basic properties of number was Frege’s criterion of success for a definition of number.

Section 56 closes with the complaint that the definitions have only fixed the sense of the phrases ‘the number 0 belongs to’, ‘the number 1 belongs to’ but they do not allow us to pick out 0 and 1 as self-subsistent objects. Dummett sees §56 as intending to address the radical adjectivalist, and sees this remark as again incomprehensibly question begging. However, if the definitions were intended to introduce numerical terms, as the use of the definite article clearly suggests, then there is nothing question begging here. Frege is simply pointing out that the phrase ‘the number 1’ has not been defined other than as part of ‘the number 1 belongs to’. That is, although the definition makes clear the contribution of the phrase ‘the number 1 belongs to’ to sentences containing it, it does not show what the contribution of ‘the number 1’ is. Since, on the interpretation defended here, Frege is attempting in §55 to define numerical terms, this points to what is obviously a significant flaw in the definitions.

In §57 Frege does finally consider adjectival uses of number. This section does not merely continue the attack on the radical adjectival definition of number, as Dummett insists, but is in fact, the first place where purely adjectival uses of number are considered. Frege begins, however, by discussing the fundamental thought –“the content of a statement of number is an assertion about a concept”¹⁹. Frege is concerned that this might lead one to conclude that numbers are properties of concepts. If we examine statements of the form ‘the number n belongs to the concept F ’, n appears as only part of what is predicated of F . As a mere part of the predicate we are not obliged to view numbers as properties of concepts. Frege then says “Precisely because it forms only an element of what is asserted, the individual number shows itself for what it is, a self subsistent object”. This may sound question begging, but it is not. The justification for this is not, as Dummett thinks, what was said in §56, but what comes immediately after. After this remark Frege says why he seeks to define numbers as objects. His first reason is that we often

¹⁹ FREGE, *The Foundations of Arithmetic*. Op. cit., §46.

attach the definite article to numbers. The second reason is that Frege believes that equations are best viewed as identities²⁰.

Now, if, on the one hand, Frege's goal were to show that numbers *really are* objects, then the reasons given in §57 are incredibly flimsy.²¹ I imagine it is for this reason that Dummett looks for further arguments against the radical adjectivalist in §56. Since the two, briefly stated, reasons are *clearly* insufficient to prove a strong metaphysical thesis, Dummett reads arguments for this thesis into the preceding sections. However, if, on the other hand, Frege's goal were to "arrive at a concept of number usable for the purpose of science" –as he explicitly states it is *at this very point in the text!*–, then these reasons may be sufficient to motivate a definition of numbers as objects.

It might be objected that if §§55-56 do not consider the radical adjectival definition of number, then they ought not appear under the general heading 'Every number is a self-subsistent object' as the sections from 55-61, in fact, do. We are now in a position to answer this objection. We saw that the definitions in §55 were put forward as being strongly suggested by the fundamental thought. These definitions were never meant to follow the adjectival strategy, but merely to pursue what might be suggested by the fundamental thought. Frege wants to show that even in these definitions, numbers are not properties of concepts. The inclusion of §§55-56 under this heading is to show that the fundamental thought does not force us to view numbers as properties of concepts. If Frege is right that the definitions given in §55 'suggest themselves so spontaneously in light of our previous results [the fundamental thought]', then showing that they don't commit us to a view of numbers

²⁰ Frege does give an argument for why we should view equations as identities in (FREGE, *Logic in mathematics*, Op. cit., p. 223). Here Frege argues against interpreting an equation as a problem and a solution. If we interpret the left hand side of, say, ' $4 + 3 = 7$ ' as a problem and the right hand side as a solution, then what are we to make of ' $(4 + 3) + 2 = 9$ '? It seems the left hand side would then be a problem of adding 2 to a problem. This argument may not rule out viewing equations as identities between entities of types other than objects, but shows that his interpretation of them as identities avoids a problem with an alternative view.

²¹ William Demopoulos ("Critical notice of Michael Dummett's 'Frege'". In: *Philosophy of mathematics* Canadian Journal of Philosophy, 1993. vol. 23, no. 3, p. 477-496) argues that there is a further kind of justification for Frege's taking numbers to be objects. Taking numbers to be objects allows Frege to prove that there are infinitely many numbers. Demopoulos compares this to the derivation of Kepler's laws from Newton's theory. In the second case it would be absurd to claim that the derivation of Kepler's laws was merely a motivation for Newton's theory, and that it offers no justificatory support. I admit that the derivation of the well known properties of numbers does offer some kind of support for the claim that numbers are objects. But exactly how it lends support and whether or not it lends any support to a metaphysical realist's interpretation of the claim that numbers are objects, I will not settle here.

as properties of concepts is important, since Frege intends to define numbers as objects. Dummett eventually puts forward an adjectival definition of number that avoids the problems that Frege identifies with his own definition, but does not defend the claim that these definitions are suggested immediately by the fundamental thought. I will not myself weigh in on whether Frege is right that it is the definitions in the form he provides which most naturally suggest themselves in light of the fundamental thought. But I will simply point out that *if he is right*, then the fundamental thought does not give us any reason to treat numbers as anything other than objects. This is because the definitions contain, explicitly and as I have argued intentionally, numerical terms.

Frege does eventually, still in §57, consider constructions such as ‘Jupiter has four moons’. In these constructions, ‘four’ is not a mere part of a predicate. Here, for the first time in these sections, a straightforwardly adjectival construction is under consideration. But we see that Frege makes no attempt to show that interpreting numbers as objects gets things right in some ultimate metaphysical sense, and that interpreting numbers adjectivally would get things wrong. His attitude at this point is purely pragmatic. He merely claims (as quoted earlier): “we should not (...) be deterred by the fact that in the language of everyday life number appears also in attributive constructions. That can always be got round”²².

Remember that for the present purposes, ‘metaphysical realism’ was defined in terms of the criteria one accepts for acceptability of the account one is putting forward. That is, does one take an important criteria to be that the account reflects how things actually stand in the world? The pragmatic attitude expressed here is incompatible with metaphysical realism. If there really is a fact of the matter of whether numbers are objects, and the measure of an account is that it reflects how things really are in the world, then this is a serious question that would deserve serious attention. Frege essentially devotes no attention to this question. All he does is give a couple of reasons for why he takes numbers to be objects. If one assigns Frege the lofty goal of showing that numbers are objects in the sense of the metaphysical realist, then Frege’s arguments must be seen as unbelievably weak. However, if one interprets Frege as having the more modest goal of providing a firm foundation for mathematics, such that numbers have the various

²² FREGE, *The Foundations of Arithmetic*, Op. cit.

features we standardly take them to have, then one can provide a far more palatable interpretation of these sections.

3. INDEPENDENCE

I have given an interpretation of Frege such that he is not defending any metaphysical position. When he claims that numbers are self-subsistent objects or that arithmetical claims state objective truths, he is merely stating what he takes to be true of our ordinary conception of number. In stating the properties of our ordinary notion of number, Frege is not engaging in metaphysics, but identifying what features a systematic treatment of number should preserve. When he does offer a systematic account, he takes it to be justified in terms of its agreement with ordinary notions: "The reader will recognize that my basic principles at no point lead to consequences that he is not himself forced to acknowledge as correct"²³.

Consider also this passage comparing psychological versus logical foundations:

It is *prima facie* improbable that such a structure could be erected on a base that is uncertain or defective. Anyone who holds other convictions has only to try to erect a similar structure upon them, and I think he will perceive that it does not work, or at least does not work so well. As a refutation in this I can only recognize someone's actually demonstrating either that a better, more durable edifice can be erected upon other fundamental convictions, or else that my principles lead to manifestly false conclusions²⁴.

What is important about this quote, for our purposes, is what he says about what he would take as a refutation of his system. So long as it does not lead to false consequences or turns out to be demonstrably inferior to another system, it should be accepted. This hardly seems like the standards of someone who thinks the primary goal of an account is to correctly reflect how things are in the world in some ultimate sense.

²³ FREGE, *The Basic Laws of Arithmetic: Exposition of the System*. Op. cit., p. 9.

²⁴ *Ibid.*, p. 25.

But what of Frege's insistence that the truth of mathematical statements is independent of anything any human does? It may seem that this independence is in conflict with the interpretation provided here, but that is not the case. We will examine this on two fronts. First, we will look at the question of the independent existence of mathematical objects. We will then turn to the question of the independence of mathematical propositions.

Let us then begin with the question of the independence existence of mathematical objects. We saw above that when Frege identifies numbers with extensions, no effort is made to show that this is true in any absolute sense. He does not hold that numbers really are and have always been extensions, but thinks that such an identification in a systematic treatment of number is justifiable in that it satisfies the two desiderata discussed in the first section. But what of other objects Frege wished to include in his systematic treatment of arithmetic? There are two of these: truth-values and courses-of-values. We see that in both of these cases Frege cites practical reasons for their introduction. Concerning courses-of-values he writes: "The introduction of courses-of-values of functions is a vital advance, thanks to which we gain far greater flexibility"²⁵. And concerning the truth-values he claims: "How much simpler and sharper everything becomes by the introduction of truth-values, only detailed acquaintance with this book can show. These advantages alone put a great weight in the balance in favor of my own conception, which indeed may seem strange at first sight"²⁶.

These hardly amount to arguments that truth-values and courses-of-values are objects in some ultimate metaphysical sense.

Of course, Frege does claim, repeatedly, that we cannot create objects with definitions. In *Grundgesetze* Frege says that the mathematician can no more create an object by definition than can the geographer create a sea by giving a particular portion of the Earth's surface the name 'the Yellow Sea'. Is Frege here making a metaphysical claim that numbers exist in some absolute sense? No, Frege is not engaging in metaphysics here at all. This can be seen from the clearly *logical* point that he goes on to make:

²⁵ *Ibid.*, p. 6.

²⁶ *Ibid.*, p. 7.

Now suppose one defines, for instance, the number zero, by saying: it is something which yields one when added to one. In so doing one has defined a concept, by specifying what property an object must have in order to fall under the concept. But this property is not a property of the concept defined. People frequently seem to fancy that by the definition something has been created that yields one when added to one. A great delusion! The concept defined does not possess this property, nor is the definition any guarantee that the concept is realized –a matter requiring separate investigation. Only when we have proved that there exists at least and at most one object with the required property are we in a position to invest this object with the proper name “zero”. To create zero is consequently impossible²⁷.

So, basic principles that introduce objects of certain kinds (e.g. courses-of-values) are to be accepted or not on practical grounds. They are to be accepted if their introduction gives us the kind of system that we are looking for –that is, in this case, one that could play the required role in the foundations of mathematics. Frege’s point, about it being impossible to create an object with a definition, is a logical and not metaphysical point. It is aimed at those who mistakenly believe that by defining a concept we thereby define into existence an object of which it is true. In neither of these cases, then, are Frege’s concerns metaphysical.

Let us then turn to the question of the independence of assertions involving number. Here again, Frege believes that it is a feature of our ordinary conception of number that ought to be preserved by a systematic account. There is nothing all that metaphysical about imagining a world without people. If a biologist claims that if there had been no humans the woolly mammoth would still exist, this is not a metaphysical claim. It does not seem any more metaphysical to claim that it is true according to our ordinary conception of number that, even in a world without people, if a certain area contains two pebbles and also two other pebbles, then there are at least four pebbles in that area. Now if the truth conditions for statements of number involved ideas in people’s minds, then nothing would be true of number in a world with no people. So Frege takes it that a definition of number should not mention people or their ideas. In a passage clearly related to the one discussed above, concerning the Yellow Sea, Frege writes: “If we say that “the North Sea is 10,000 square miles in extent” then neither

²⁷ *Ibid.*, p. 11-12.

by “North Sea” nor by “10,000” do we refer to the state of or process in our minds: on the contrary, we assert something objective, which is independent of our ideas and everything of the sort”²⁸.

Again, it is evident here that his target is psychologistic accounts of number. He is saying that the claim “the North Sea is 10,000 square miles in extent” makes no reference to people and so should be true whether or not there are any people. This independence is again identified as a desideratum of a systematic account of number. This desideratum is obviously not satisfied by a psychologistic account of number. But Frege does not hold that our presystematic understanding of number already have perfectly clear truth conditions: “Since it is only in the context of a proposition that words have meaning, our problem becomes this: To define the sense of a proposition in which number words occur. That, obviously, leaves us still a wide choice”²⁹.

Our ordinary conception of number has the feature that the truth of such claims is independent of us and any of our activities. However, in giving a systematic treatment of number, we still have some freedom in how we define number. What we want of such a definition is that it satisfies our quite practical desiderata.

Thus we see, Frege’s talk of independence does not commit him to metaphysical positions. Concerning basic principles that introduce entities, Frege sees their acceptance as being a pragmatic question. In terms of the independent truth of mathematical statements, this need not be interpreted as involving any metaphysics³⁰.

CONCLUSIONS

We saw that Frege describes the purpose of his project as arriving at ‘a concept of number usable for the purpose of science’. To this end, two

²⁸ FREGE, *The Foundations of Arithmetic*. Op. cit., §26.

²⁹ *Ibid.*, §62.

³⁰ Burge, T. (*Truth Thought Reason: essays on Frege*. Oxford & New York: Oxford University Press, 2005) repeatedly defends the position that the pragmatic aspect that I discuss are present in Frege, but at the end of the day, Frege takes the pragmatic success of his system to be evidence for its metaphysical truth. Although, I cannot, at this point, discuss Burge’s arguments in detail, I do not think he has provided sufficient argument for the claim that despite all the pragmatic elements in Frege’s writings, he is nonetheless a metaphysical realist. Burge’s main line of argument for his position is the independence just discussed. But if this need not be interpreted metaphysically, then I see little reason for a metaphysical interpretation of Frege’s writings.

desiderata were identified: the desideratum of standardness and the desideratum of harmlessness. The desideratum of standardness holds that the properties we ordinarily associate with number should also hold of numbers in a systematic treatment. The desideratum of harmlessness holds that any new properties a definition confers on number should not significantly change the practice of mathematics. A systematic treatment that satisfies these desiderata would seem to fully satisfy Frege. There is no further requirement that the systematic treatment reflect how things are in the world in some ultimate sense.

We then looked in detail at §§55-61 of *Grundlagen*. Whereas Dummett sees Frege as addressing the radical adjectivalist in §55 and §56, we saw there was nothing in those sections to suggest that. In fact, it is quite clear that Frege intends in these sections to define numerical terms. It is a further advantage of my interpretation, unlike that of Dummett, that it can make sense of the Julius Caesar objection as it appears in these passages. Only if it is numerical terms being defined is Julius Caesar a legitimate substitution instance. Finally, in §57, when Frege does consider purely adjectival uses of number, he expresses a purely pragmatic attitude. This is not consistent with a *metaphysical view* of numbers as objects.

In the last sections we looked at Frege remarks concerning the independent existence of number and the independent truth of arithmetical propositions. It was shown that on neither of these issues need we interpret Frege as holding a metaphysical thesis. When he says we cannot create objects with our definition, he is making a logical, rather than metaphysical, point. When he claims that arithmetical propositions are true completely independently of us, he is saying that we ought not define number such that truths of number depend on truths about us.

REFERENCES

- BEANEY, M. (1996). *Frege: Making Sense*. London: Duckworth.
- _____. (2004). "Carnap's conception of explication: from Frege to Husserl?" *In*: AWODEY, S. & KLEIN, C. (Eds.), *Carnap Brought Home: The View from Jena* (pp.117-150). Chicago and La Salle, IL: Open Court.
- BURGE, T. (2005). *Truth Thought Reason: essays on Frege*. Oxford & New York: Oxford University Press.
- DEMOPOULOS, W. (1993). "Critical notice of Michael Dummett's 'Frege'". *In*: *Philosophy of mathematics' Canadian Journal of Philosophy*, vol. 23, no. 3, p. 477-496.
- DUMMETT, M. (1991). *Frege: Philosophy of Mathematics*. Cambridge, MA: Harvard University Press.
- _____. (1993). "Frege as a realist". *In*: SLUGA, H. (Ed.), *Meaning and Ontology in Frege's Philosophy*, Vol. 3 of *The Philosophy of Frege*. (pp. 109-122). New York & London: Garland.
- FREGE, G. (1967). *The Basic Laws of Arithmetic: Exposition of the System*. Edited and translated by Montgomery Furth. Berkeley and Los Angeles: University of California Press.
- _____. (1979). "Logic in mathematics". *In*: HANS HERMES, F. K. & KAMBARTEL, F. (Ed.). *Posthumous Writings*. (pp. 203-250). Oxford: Basil Blackwell.
- _____. (1980). *The Foundations of Arithmetic*. 2d. ed. Evanston, IL: Northwestern University Press.
- _____. (1984a). "Review: Husserl, philosophy of arithmetic". *In*: McGUINNESS, B. (Ed.). *Collected Papers on Mathematics, Logic and Philosophy*. (pp. 293-340). Oxford: Basil Blackwell.
- _____. (1984b). "On the foundations of geometry: second series". *In*: McGUINNESS, B. (Ed.). *Collected Papers on Mathematics, Logic and Philosophy*. (pp. 293-340). Oxford: Basil Blackwell.
- RECK, E. H. (1997). "Frege's influence on Wittgenstein: reversing metaphysics versus the context principle". *In*: TAIT, W. W. (Ed.). *Early Analytic Philosophy: Frege, Russell, Wittgenstein*. (pp. 123-85). Chicago and La Salle, IL: Open Court.
- _____. (2007). "Frege on truth, judgement, and objectivity". *In*: *Grazer Philosophische Studien*, vol. 75, no. 1, p. 149-173.
- RICKETTS, T. (1986). "Objectivity and objecthood: Frege's metaphysics of judgement". *In*: HAAPARANTA, L. & HINTIKKA, J. (Eds.). *Frege Synthesized*. (pp. 65-95). Dordrecht: Reidel.
- _____. (1996). "Frege on logic and truth". *In*: *Proceedings of the Aristotelian Society Supplement*, vol. 70, p. 121-140.

WEINER, J. (2007). "What's in a numeral? Frege's answer" *In: Mind*, vol. 116, no. 463, p. 677-716.

WRAY, K. B. (1995). "Reinterpreting section 56 of Frege's *The Foundations of Arithmetic*". *In: Auslegung: A Journal of Philosophy*, vol. 20, no. 2, p. 76-82.

THE LIAR AND THE NEW T-SCHEMA

EL MENTIROSO Y EL NUEVO ESQUEMA T

STEPHEN READ

University of St Andrews, United Kingdom. slr@st-andrews.ac.uk

RECIBIDO EL 30 DE SEPTIEMBRE DE 2010 Y APROBADO EL 25 DE OCTUBRE DE 2010

RESUMEN

Desde que Tarski publicó su estudio sobre el concepto de verdad en los años 30, ha sido una práctica ortodoxa el considerar que toda instancia del esquema T es verdadera. Sin embargo, algunas instancias del esquema son falsas. Éstas incluyen las instancias paradójicas ejemplificadas por la oración del mentiroso. Aquí se demuestra que un esquema mejor permite un tratamiento uniforme de la verdad en el que las paradojas semánticas resultan ser simplemente falsas.

PALABRAS CLAVE

Bradwardine, Carnap, Davidson, esquema T, paradoja del mentiroso, Tarski, teoría de la verdad.

ABSTRACT

Since Tarski published his study of the concept of truth in the 1930s, it has been orthodox practice to suppose that every instance of the T-schema is true. However, some instances of the schema are false. These include the paradoxical instances exemplified by the Liar sentence. It is shown that a better schema allows a uniform treatment of truth in which the semantic paradoxes turn out to be simply false.

KEY WORDS

Bradwardine, Carnap, Davidson, T-scheme, liar paradox, Tarski, truth-theory.

1. TARSKI'S T-SCHEME

Carnap, in his *Autobiography*¹ tells an anecdote about a meeting with Tarski in Vienna in the early 1930s. Both Carnap and Tarski had been struggling with the problem of providing a satisfactory definition of truth. Carnap asked Tarski to explain the basic idea: how can one define the truth of an empirical sentence like 'This table is black'. "'This table is black' is true," replied Tarski, "if and only if this table is black". Coffa recalls that Carnap later described this as the moment when "the scales fell from my eyes"².

Over the years since Tarski published his study of the concept of truth³, it has been an unquestioned orthodoxy that every instance of the T-scheme:

x is true if and only if p (T)

is true, where what replaces ' x ' is a name of a sentence whose translation into the metalanguage replaces ' p '. Indeed, in the homophonic case of identity (in sound or in fact) of object language and metalanguage, what replaces ' x ' is a quotation-name of the sentence which replaces ' p '.

Nonetheless, not all instances of (T) are true, even where the very same sentence replaces ' x ' and ' p ', the first in quotation-marks. For example,

'I am tired' is true if and only if I am tired (1)

and

'That book was stolen' is true if and only if that book was stolen (2)

are not true unless the person who utters (1) is the same as the utterer of the sentence cited, and unless the book indicated by the demonstrative pronoun in (2) is the very same as that demonstrated in the sentence whose truth is at stake, as Davidson⁴ pointed out. Moreover, even in

¹ CARNAP, R. "Intellectual autobiography in P. Schilpp". In: *The Philosophy of Rudolf Carnap*. (pp. 3-84). La Salle, IL: Open Court, 1963. p. 60.

² COFFA, J. A. *The Semantic Tradition from Kant to Carnap: To the Vienna Station*. Cambridge: Cambridge University Press, 1991. p. 304.

³ Cfr. TARSKI, A. (1936). "The establishment of a scientific semantics". In: WOODGER, J. (Ed.). *Logic, Semantics, Metamathematics*, 2d Ed. Corcoran, J. (Tr.). Indianapolis: Hackett, 1983. p. 152.

⁴ Cfr. DAVIDSON, D. "True to the facts". In: *Journal of Philosophy*, 1969. vol. 66, p. 304. Reprinted in his *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press, 1984. p. 37-54.

the absence of indexicals, the validity of the T-scheme is not assured. Hintikka's example⁵ was

'Any man is mortal' is true if and only if any man is mortal (3)

'Any' always takes widest scope, so that whereas 'any' in the quoted sentence acts as a universal quantifier within it, the occurrence of 'any' on the right-hand side functions as a universal quantifier with wide scope over the whole of (3), and so as an existential quantifier with scope restricted to the right-hand side. Thus homophonic translation fails.

Despite such counterexamples, (T) are often taken as a truism. For example, in his paper 'The Establishment of a Scientific Semantics', Tarski wrote:

Quite generally, we shall accept as valid every instance of the form

the sentence x is true if and only if p

where 'p' is to be replaced by any sentence of the language under investigation and 'x' by any individual name of that sentence⁶.

More recently, Paul Horwich wrote:

for any declarative sentence 'p' our language generates an equivalent sentence 'The proposition that p is true'⁷.

and

(...) the concept of truth is entirely captured by stipulating the equivalence schema, 'The proposition that p is true if and only if p' -where p can be replaced by any declarative sentence⁸.

⁵ HINTIKKA, J. "A counterexample to Tarski-type truth-definitions as applied to natural languages". In: *Philosophia*, 1975. vol. 5, p. 207.

⁶ TARSKI. *The establishment of a scientific semantics*. Op. cit., p. 404.

⁷ HORWICH, P. *Truth*. 2d. Ed. Oxford: Clarendon Press, 1998. p. 4-5.

⁸ HORWICH, P. "In the truth domain". In: *Times Literary Supplement*, 4711, 1993. p. 28.

Lastly, Aladdin Yaqub writes:

The concept of truth is *completely* and *correctly* defined by the Tarskian scheme, whose instances are *all* the biconditionals obtained from the phrase '*x* is true if and only if *p*' by substituting any sentence for '*p*' and any expression which stands for that sentence for '*x*'⁹.

Yaqub invokes Tarski, Ayer, Horwich and Gupta in support of his claim that "*the Tarskian schema is the most fundamental intuition about the concept of truth*"¹⁰.

Tarski did not propose (T) as a definition of truth, though others, e.g., Horwich, have done so since. They all describe (T) as a truism, as something so obviously correct as to need no argument, at least, provided the obvious counterinstances, such as (1), (2) and (3) are excluded. Tarski himself presented (T) as a condition of what he called "material adequacy". It was a necessary condition on any adequate definition of truth that it generates all instances of (T) as theorems:

A formally correct definition of the symbol '*Tr*' (...) will be called an adequate definition of truth if it has the following consequence:

all sentences which are obtained from the expression '*x* ∈ *Tr* if and only if *p*' by substituting for the symbol '*x*' a structural-descriptive name of any sentence of the language in question and for the symbol '*p*' the *expression which forms the translation of this sentence into the metalanguage*¹¹.

Tarski's account was for many years described as a version of the correspondence theory of truth –an acceptable version of a theory relieved of its suspect ontology of facts and correspondence. The (T)-scheme presents in a stark and nominalistically acceptable form the correlation of language, on the left-hand side, and world on the right-hand side. What replaces '*x*' refers to a sentence or other truth-bearer, and what replaces '*p*' refers to, or describes, how things are. Thus the (T)-scheme can be summed up as expressing the idea that a sentence is true if things are as it says they are:

⁹ YAQUB, A. *The Liar Speaks the Truth*. Oxford: Oxford University Press, 1993. p. 3.

¹⁰ *Ibid.*, p. 27.

¹¹ TARSKI. The establishment of a scientific semantics. *Op. cit.*, p. 187-188.

x is true if and only if things are as x says they are (CP)

Crispin Wright describes (CP) as “*the correspondence platitude*”¹², and Julian Dodd agrees: “*a platitude it certainly is (...) it says nothing that would be disagreed with by a correspondence theorist, coherence theorist or deflationist*”¹³.

In this paper, I want to show that (CP) is at best equivocal, and (T) is false. This is not just a matter of showing that not all instances of (T) are true. Many proponents of (T) admit that. They nonetheless claim that (T) offers a good account (partial or complete) of the truth-conditions of those sentences for which it is true. However, this is not so. Where it is true, (T) is true by lucky chance. Where true, its instances are also instances of a better scheme, the right scheme of truth. (T) is very far from being a truism, or a good test of the adequacy of a truth-theory.

2. HOW THINGS ARE

How can things be as x says they are unless x is true? Clearly, ‘snow is white’ says that snow is white, and ‘snow is white’ is true just when snow is as ‘snow is white’ says it is, that is, when snow is white. But ‘Every even number greater than 2 is the sum of two primes’ not only says that every even number greater than 2 is the sum of two primes, but also in particular that, e.g., 4 is the sum of two primes. So 4’s being the sum of two primes is as Goldbach’s Conjecture says things are. But Goldbach’s Conjecture is not true just because 4 is the sum of two primes. 4’s being the sum of two primes is necessary for the truth of Goldbach’s Conjecture, but it is not sufficient. So (CP) needs more careful expression. What was meant was

x is true if and only if things are wholly as x says they are, (S)

or

x is true if and only if however x says things are, they are (S’)

A sentence says many things, and however it says they are, they must be, for the sentence to be true. Anything of the form ‘ p and q ’ says *inter alia* that p , but the truth of p does not suffice for that of ‘ p and q ’.

¹² WRIGHT, C. J. G. *Truth and Objectivity*. Cambridge, MA: Harvard University Press, 1992. p. 27.

¹³ DODD, J. “Is truth supervenient on being?” In: *Proceedings of the Aristotelian Society*, 2001-2002. vol. 102, p. 75.

Indeed, the phrase ‘how x says they are’ or even ‘however x says they are’ is imprecise in another way. ‘All men are mortal’ says that if Socrates is a man then Socrates is mortal. But Socrates was a man. So if ‘All men are mortal’ is true, Socrates was mortal. Whether the universal sentence actually says that Socrates was mortal is, perhaps, unclear. If Socrates was not mortal, then ‘All men are mortal’ is false –unless Socrates was not a man after all. If x is true, things must be however x implies they are. Conversely, if things are however x implies they are, x will be true. “Safety first” would support the identification of ‘however x says they are’ with ‘what x implies’. This may make ‘how x says they are’ more precise than is natural, but in the context of (CP) it is welcome precision. We find such a conception of ‘saying that’ as closed under implication in Wittgenstein’s *Tractatus*: “A proposition affirms every proposition that follows from it (...) If p follows from q , the sense of ‘ p ’ is contained in the sense of ‘ q ’”¹⁴. By ‘follows from’, Wittgenstein here clearly intended the logical truth of the conditional (as John Corcoran pointed out to me). Elementary propositions are logically independent and p follows from q only if the truth-conditions or truth-grounds (Sect. 5.101), of q contain those of p . For Wittgenstein, the conditional itself was a material one. I will follow him in this, for the purposes of this paper. Thus logical implication is presently strict implication, that is, the necessitation (or logical truth) of the material conditional. Certain move in the argument to follow work only for such an implication.

Let us represent ‘ x says that p ’ by ‘ $x:p$ ’, where ‘ x ’ designates a sentence, strict implication by ‘ \Rightarrow ’ and the material conditional by ‘ \rightarrow ’. Then we can represent the closure of ‘saying that’ under logical implication by

$$(\forall p, q)((p \Rightarrow q) \rightarrow (x:p \Rightarrow x:q)) \text{ (K)}$$

The context $x:p$ is an intensional context for p , allowing substitution only of logical equivalents (and then, only if we accept the closure of $x:p$ (saying that) under ‘ \Rightarrow ’)¹⁵. The truth-condition given by (S) or (S’) is then represented by

¹⁴ WITTGENSTEIN, L. *Tractatus Logico-Philosophicus*. Pears, D. & McGuinness, B. (Trs.). London: Routledge and Kegan Paul, 1961. Sects. 5.122 and 5.124.

¹⁵ Indeed, if ‘ \Rightarrow ’ were ‘ \rightarrow ’ and x said anything, say p , then since $(p \rightarrow q) \vee (p \rightarrow \neg q)$ is a tautology, we would have for all q , $x:q \vee x:\neg q$, which would trivialize the whole project and render $x:p$ extensional. What one would like for ‘ \Rightarrow ’ is some enthymematic conditional, which includes the case of Socrates above without trivializing the whole notion. But identifying such a conditional is, presently, I believe, an unrealized project.

$$Tx \Leftrightarrow \forall p(x:p \rightarrow p) \text{ (A)}$$

Here, ' \Leftrightarrow ' represents strict equivalence. For Tx and $\forall p(x:p \rightarrow p)$ do not just happen to be equivalent. The right-hand side of (S) gives a philosophical analysis, or explication, or truth. (A) contrasts with (T) in two important ways. First, (A) is an analysis of truth; (T) is merely, for Tarski, a condition of material adequacy, a test of the accuracy of his and others' theories. In particular, (A) is not the final analysis of truth: that awaits an analysis of 'saying that'. Thus (A) need not conflict with Tarski's requirement that all semantic notions be defined and eliminated, any more than does his own account of truth:

$$Tx \Leftrightarrow \forall s(s \text{ is a sequence} \rightarrow s \text{ satisfies } x).$$

Tarski defined truth in terms of satisfaction, but then proceeded to define satisfaction in non-semantic terms.

Secondly, (A) is a logical equivalence, whereas (T) was for Tarski merely a material equivalence. For Horwich, the partial definitions whose form is (T) are given as material equivalences, but, he says, their a priori role allows us to derive the corresponding strict equivalences¹⁶. That is fitting for his use of (T) in defining truth, a use which Tarski did not make. Accordingly, if (A) is a logical equivalence in providing the correct analysis of truth, we must assume that what a sentence says, it says of necessity too. Of course, any particular form of words expresses what it does only contingently. When I speak of the truth of sentences, and what they say, in what follows, I will be speaking of a sentence taken as meaning what it does – of a proposition, in Horwich's terminology, or of a sentence in a context of speaker, place and time.

If we take the more cautious line of rejecting (K), we can still capture (S) and (S') by

$$Tx \Leftrightarrow \forall p(\exists q(x:q \wedge (q \Rightarrow p)) \rightarrow p)^{17} \text{ (A')}$$

Whereas (A) expresses the thought that a sentence is true if things are however it says they are, (A') spells out that it is true if things are

¹⁶ Cfr. HORWICH, P. *Truth*. 2d. Ed. Oxford: Clarendon Press, 1998. Chap. 2, Sect. 3.

¹⁷ Note that (K), (A) and (A') should be thought of as expressions in type theory, where x has type i (of individuals) and p has type o (of propositions). See: e.g., CHURCH, A. "A formulation of the simple theory of types". *In: Journal of Symbolic Logic*, 1940. vol. 5, p. 56-68.

however what it says (logically) implies them to be. They are equivalent just if we identify what a sentence says with what it implies. Of course, taking ' \rightarrow ' to be material implication, (A) and (A') will make true all sentences which say nothing. One solution is to conjoin $(\exists p) x:p$ to the right-hand side of (A) and (A'). But for simplicity, I will assume in what follows that each sentence to which we apply (A) and (A') does say something.

Another possible objection to (A) –and to (S)– concerns ambiguity. A different way for a sentence to say more than one thing is for it to be ambiguous. But if x has two meanings, e.g., 'Visiting relatives can be boring', we do not require that both senses be the case for x to be true –either will suffice.

In reply, remember that the point of formalization is to avoid ambiguity. Accordingly, we must remove ambiguity from x and require that all sentences governed by (A) be unambiguous. For example, we might deem 'Visiting relatives can be boring' to be a pair of homophones, as we do with 'I have a pair/pear'. Nonetheless, an unambiguous sentence, like 'All men are mortal', can say many things –that Socrates is mortal, that Plato is mortal and so on. To be true, all the things it says must hold– things must be however it says they are.

The crucial advantage of (A), or (A'), over (CP) is that it makes clear that everything which a sentence says or implies must be the case for the sentence to be true. (CP) is equivocal: x may say things are in many ways. They are all severally necessary, but only jointly sufficient. Things must be not only how x says they are, but however x says they are, in order that x be true.

Tarski's later discussion helps us see where he went wrong. Denoting 'snow is white' by ' S ', he writes:

What do we mean by saying that S is true...? The answer to this question is simple: (...) by saying that S is true we mean simply that snow is white (...) By eliminating the symbol ' S ' we arrive at the following formulation:

(1) 'snow is white' is true if and only if snow is white¹⁸.

¹⁸ TARSKI. The establishment of a scientific semantics. Op. cit., p. 64.

But he generalizes this later on the same page:

Partial definitions of truth analogous to (1) (...) can be constructed for other sentences as well. Each of these definitions has the form:

(3) '*p*' is true if and only if *p*,

where '*p*' is to be replaced on both sides of (3) by the sentence for which the definition is constructed¹⁹.

That snow is white is, arguably, all that 'snow is white' says. So the right-hand side of (1) gives a sufficient, as well as necessary condition for its truth. But not all cases are as straightforward as this. In general we need to do more than simply repeat '*p*' on the right-hand side. (T) is a special case of (A) where, as for 'snow is white'; a sentence reveals explicitly all that it says. The crucial importance of using (A), and the disastrous consequences of using (T) in its place, will be demonstrated in Sects. 3 and 4.

3. THE "LIAR PARADOX"

Tarski and others exclude the Liar paradox from the scope of the T-scheme. Tarski did it by forbidding semantic closure, and so making the semantic paradoxes inexpressible²⁰. Horwich²¹ simply excludes the paradoxical instances of (T) from the infinite conjunction of its instances which he takes to define truth. The consequence is that neither author can offer an account of the semantics of the Liar paradox and its ilk. For the one, the material test of the correctness of his theory of truth is that the right instances of (T) be derivable in the theory; for the other, the theory of truth is constituted by the conjunction of the correct instances. Each implicitly concedes, therefore, that not all instances of (T), even the homophonic ones where what replaces '*x*' is a quotation-name of the sentence which replaces '*p*', are true. Nonetheless, they persevere in keeping (T) as the cornerstone of their theory of truth, rather than looking for a better scheme. A better scheme would be one all of whose instances are true, one which did not need an ad hoc exclusion clause, excluding any counterexamples by *fiat*.

¹⁹ *Ibidem*.

²⁰ See TARSKI. Sect. 1 and p. 402.

²¹ See HORWICH. Sect. 10.

There are many ways to formulate the Liar paradox, but the general idea is to find a sentence-name L such that L and ' L is not true' are interchangeable in light of some fact about L (e.g., where L is written, or through the correlations of some system of Gödel-numbering, and so on). Recall that we are assuming that a sentence (or proposition) necessarily says what it says, even, for example, in the case of contingent paradox²². Then since L implies itself, it implies that L is not true (by substitution), and indeed, if ' L says that' is closed under implication (as proposed in Sect. 2) or if the connection is close enough, L says that L is not true. Thus we can substitute ' L is not true' for ' p ' and ' L ' for ' x ' in (T), or (CP), to obtain the paradoxical instance:

L is true if and only if L is not true (P)

Without serious and damaging revisions to logical principles, (P) is false, for it entails:

L is true and L is not true (P')

Those who reject Ex Falso Quodlibet (the principle that a contradiction implies every sentence) may attempt to maintain that (P') is true without collapsing into triviality. But the Curry paradox, finding for an arbitrary sentence q , a sentence Cq which is interchangeable with ' Cq is true then q ', produces the paradox of triviality without recourse to contradiction. Substituting ' Cq ' for ' x ' in (T), we obtain

Cq is true iff if Cq is true then q ,

and by familiar steps of Modus Ponens and Contraction, we get a proof of q . Since q was arbitrary, the truth-theory based on (T), and (CP), is shown to be unacceptable. (T) is simply false, for it abounds in false instances.

4. TRUTH-THEORY WITHOUT PARADOX

What we have recalled in Sect. 3 is that if we look to (T) for an account of the semantics of L and Cq , the result will be trivial and paradoxical—in a word, wrong. The mistake encapsulated in (T) is to suppose that the expression replacing ' p ' on the right-hand side fully expresses

²² See KRIPKE, S. "Outline of the theory of truth". In: *Journal of Philosophy*, 1975. vol. 72, p. 691.

its own truth-condition. This may seem a plausible assumption. But at least in the case of the semantic paradoxes, as we will see, it is a mistaken assumption. For sentences imply other things which they do not themselves explicitly say. By the lights of (A), they say them nonetheless. By the lights of (A) or (A'), if those other things are not the case, then what the original sentence says cannot hold either, and so the original sentence is false. Reading (CP) as (T) misses that implication. (T) Identifies the truth of a sentence with things being how it says they are; but not with however it says they are. Reading (CP) as (S), and so as (A) or (A'), includes in the truth-conditions of a sentence all its consequences.

Consider *L*, for example. *L* says that *L* is not true, that is, $L: \neg TL$. It may say more: call it *q*, that is, supposes $L: (\neg TL \wedge q)$. (If $\neg TL$ is all that *L* says, we can simply let *q* be $\neg TL$, or a tautology.) By (A),

$$TL \Leftrightarrow \forall p(L:p \rightarrow p)$$

i.e.,

$$TL \Leftrightarrow (\neg TL \wedge q)$$

since $\neg TL \wedge q$ is all *L* says, whence

$$\neg TL \Rightarrow \neg (\neg TL \wedge q)$$

i.e.,

$$\neg TL \Rightarrow (TL \vee \neg q).$$

If saying that is closed under ' \Rightarrow ', as (K) says, then, since $L: \neg TL$,

$$L: (TL \vee \neg q)$$

and so (since $L:q$)

$$L: ((TL \vee \neg q) \wedge q)$$

whence

$$L: TL$$

since $((TL \vee \neg q) \wedge q) \Rightarrow TL$. Thus, if *L* says $\neg TL$, it says *TL* as well. That is, $L: (\neg TL \text{ c } TL)$. So by (A),

$$TL \Leftrightarrow (\neg TL \wedge TL \wedge \dots)$$

whence

$$\neg TL.$$

The argument can be repeated using (A') in place of (A), and dropping the assumption that saying that is closed under '⇒'. Assume again that $L: (\neg TL \wedge q)$. By (A'), since $\neg TL \wedge q$ is all that L says,

$$TL \Leftrightarrow \forall p(((\neg TL \wedge q) \Rightarrow p) \rightarrow p),$$

so

$$\neg TL \Rightarrow \exists p(((\neg TL \wedge q) \Rightarrow p) \wedge \neg p),$$

i.e.,

$$\neg TL \Rightarrow \exists p(\neg p \wedge (\neg p \Rightarrow (TL \vee \neg q))),$$

so

$$\neg TL \Rightarrow \exists p(\neg p \wedge (TL \vee \neg q)),$$

whence

$$\neg TL \Rightarrow (\neg TL \rightarrow \neg q),$$

which by Contraction yields,

$$\neg TL \Rightarrow \neg q.$$

Contraposing, we have

$$q \Rightarrow TL.$$

Using (A') once again,

$$TL \Leftrightarrow (\neg TL \wedge TL \wedge \dots)$$

whence

$$\neg TL.$$

It may not be immediately obvious what a breakthrough this is. Of course L is not true. By (A)

$$TL \Leftrightarrow (L: \neg TL \rightarrow \neg TL) \wedge \dots$$

whence

$$TL \Rightarrow (L: \neg TL \rightarrow \neg TL).$$

Since $L: \neg TL$,

$$TL \Rightarrow \neg TL$$

and so

$$\neg TL.$$

But the longer argument concerning L shows more than that L is not true. In Sect. 3, using the (T)-scheme, we concluded not only that L was not true but also that L was true, and so landed in paradox. This reasoning fails when the truth-conditions of L (and all sentences) are given by (A), or (A'). To use (T) to infer that L is true, all we need to show is that L is not true, as in Sect. 3. To use (A) or (A') to show that L is true we need to show both that L is not true and, in addition, that L is true. But that is impossible. L cannot be true, for to be true, it would have to be both true and not true. Nothing can be both true and not true. So L cannot be true. L is not only not true, it cannot (also) be true, as (A) reveals.

Note that I have not shown that $\neg TL \Rightarrow TL$. The closure principle (K) implies that L says whatever is implied by what L says, and, by the argument above, it follows from the fact that $L: \neg TL$ that $L: TL$. But $\neg TL \Rightarrow TL$ is in fact false. It entails $\neg TL \rightarrow TL$, which is equivalent to TL . But TL is false, as we have seen, and consequently, so is $\neg TL \Rightarrow TL$.

The problem with (T), as a condition of truth, is that it implies the impossible, for it declares (unless, of course, suitably restricted) that L is both true and not true. Thus (T) is false. (T) has the false instance:

$$TL \Leftrightarrow \neg TL.(†)$$

The mistake made by Tarski, Horwich and others is to think that a suitably restricted version of (T) can serve as a determiner of the right theory of truth. It cannot. Either for no good reason other than to preserve the (universal) truth of (T), the counterinstances are deemed ill-formed, a thoroughly ad hoc manoeuvre; or the counterinstances are simply excluded from the theory of truth by *fiat*, again ad hoc and involving the further consequence that there are apparently well-formed sentences (such as *L*) for which there is no theory of truth. How, unless we can apply our theory of truth to *L* can we discover that it is paradoxical?

The solution is ready to hand. Abandon (T) and realise that the correct theory of truth is given by (A) –or (A'), depending on one's preferred theory of meaning– governing all well-formed sentences in a semantically closed language. As applied to *L*, we obtain the correct truth-condition:

$$TL \Leftrightarrow (\neg TL \wedge TL).(\dagger\dagger)$$

I said that (CP) equivocates, and perhaps (T) does so too, or at least has been misunderstood. Tarski requires that any adequate theory of truth for a language must generate all instances of (T) where what replaces '*p*' is a translation of the sentence whose name replaces '*x*'. But a proper translation must spell out all that the sentence says. So arguably, (\dagger) is not a correct instance of (T), for the right-hand side does not say all that *L* says. Rather, the correct instance of (T) is ($\dagger\dagger$), whose right-hand side says all that *L* says.

5. PARADOX WITHOUT TRUTH

Tarski²³ observed that one can repeat the Liar reasoning in the absence of a truth-predicate, simply with a quotation-functor. Even after replacing (T) by (A), there is a danger that we might similarly find paradox has returned, through use of the 'says that' predicate. For consider

$$\forall p(C:p \rightarrow \neg p)(C)$$

Clearly $C: \forall p(C:p \rightarrow \neg p)$. Suppose that $\forall p(C:p \rightarrow \neg p)$. By Universal Instantiation,

²³ TARSKI, The establishment of a scientific semantics. Op. cit., p. 161.

$$C: \forall p(C:p \rightarrow \neg p) \rightarrow \neg \forall p(C:p \rightarrow \neg p)$$

so by Modus Ponens,

$$\neg \forall p(C:p \rightarrow \neg p)$$

and so by *reductio*,

$$\neg \forall p(C:p \rightarrow \neg p),$$

whence

$$\exists p(C:p \wedge p).(*)$$

We might be tempted to infer

$$\forall p(C:p \rightarrow \neg p)$$

from (*), since that is what C seems to say. But we have already shown $\neg \forall p(C:p \rightarrow \neg p)$, and so contradiction would have returned, without the involvement of (T) or (A) and the truth-predicate.

However, although $C: \forall p(C:p \rightarrow \neg p)$, this may not be all that C says. Suppose C also says that q , that is, $C: (\forall p(C:p \rightarrow \neg p) \wedge q)$. Then by (A),

$$TC \Leftrightarrow (\forall p(C:p \rightarrow \neg p) \wedge q)$$

whence

$$(\neg \forall p(C:p \rightarrow \neg p) \vee \neg q) \Rightarrow \neg TC$$

and so

$$\neg \forall p(C:p \rightarrow \neg p) \Rightarrow \neg TC,$$

whence, since $\forall p(C:p \rightarrow \neg p) \Rightarrow \neg \forall p(C:p \rightarrow \neg p)$,

$$\forall p(C:p \rightarrow \neg p) \Rightarrow \neg TC$$

and so

$$C: \neg TC.$$

Thus it is true that $\exists p(C:p \wedge p)$, for $C:\neg TC \wedge \neg TC$. So from (*) we cannot legitimately infer that whatever C says is the case. Some of what C says is the case, namely, $\neg TC$. But (*) is no ground for inferring that whatever C says is the case. In particular, $\forall p(C:p \rightarrow \neg p)$ is not the case, since C is false.

Although this frees C from the threat of immediate paradox, it raises a further worry. For it is crucial to the present solution to the paradoxes that not every proposition say of itself that it is true. The reason that L , C and Cq turn out not to be paradoxical is that they say of themselves that they are true, like the Truth-teller; but since they also directly or indirectly say that they are not true, things cannot be as they say they are, and so by (A) they are not true but false.

But if we could show that every sentence said (albeit indirectly) of itself that it was true, the truth-condition (A) would be crippled. For (A) says that a sentence is true only if things are as it says they are, and if part of what it said was that it was true, (A) would, correctly but unhelpfully, endorse a sentence as true if it was true, but would not reveal whether it was true or not. Every sentence (bar the paradoxical ones) would become a Truth-teller, true if true and false if not, but of undetermined truth-value.

What induces this anxiety that every sentence is a truth-teller? The thought is this: implicit in endorsing (A) as the right truth-condition (in place of (T)) was that 'saying that' is closed under implication –every sentence says (implicitly, or indirectly) whatever is implied by what it says. But by (A), things being as it says they are implies that a sentence is true. It seems to follow that every sentence says that it is true. That it is true is implied by whatever it says, so if it says whatever it implies, it says that it is true.

Fortunately, this is a specious argument, which we can see from our formalization of it. (A) says

$$Tx \Leftrightarrow \forall p(x:p \rightarrow p)$$

and the closure of 'saying that' under implication, (K), is

$$(\forall p, q)((p \Rightarrow q) \rightarrow (x:p \Rightarrow x:q)).$$

These do not entail $x:Tx$. For a counterexample, read $x:p$ as $p \in X$, where X is a theory, that is, a collection closed under logical implication. Then the closure assumption is true, i.e.,

$$(\forall p, q)((p \Rightarrow q) \rightarrow (p \in X \Rightarrow q \in X)).$$

We want $x:Tx$ to be false. (A) identifies Tx with $\forall p(p \in X \rightarrow p)$. So we need to show that it is possible that $\forall p(p \in X \rightarrow p) \notin X$. What $\forall p(p \in X \rightarrow p)$ says is that X is correct –that X contains only truths. But theories can be mistaken, and it would be an arrogant theory which included the statement $\forall p(p \in X \rightarrow p)$, which says that everything X says is true. Thus $x:Tx$ does not follow from (A) together with the closure principle.

Thus the argument that every sentence says of itself that it is true is sophistical. A sentence says whatever it says, but it does not say that whatever it says is the case –it does not say that things are as it says they are. The Truth-teller says that. But that is what is wrong with the Truth-teller. To say that things are as it says they are is to say nothing. Things must be as a sentence says they are for it to be true, but it does not (in general) say that things are as it says they are.

It follows that paradox, with or without truth, is prevented by the adoption of the correct account of truth, namely, the truth-condition (A).

6. SOME HISTORY

The argument in Sect. 4 to the effect that L says that L is true since it says that L is not true is due to Thomas Bradwardine, briefly Archbishop of Canterbury in 1349 before succumbing to the Black Death. His treatment of the semantic paradoxes dates from some 25 years earlier, when he was a young teacher at the University of Oxford. His solution was taken up and adapted by John Buridan and Albert of Saxony at the University of Paris, and in Buridan's version was discussed in the 1950s and sixties by Ernest Moody (1953) and Arthur Prior (1962). However, both John and Albert omit the proof from Sect. 4., John (2001, 969) simply asserting that every sentence implies its own truth, Albert (1974, f. 43rb) inferring that every sentence says that it itself is true from the assumptions that every (affirmative) sentence says that its subject and predicate are identical, and that every such sentence whose subject and predicate are identical is true (and likewise for negatives). It was this reasoning of Albert's which inspired the sophistical reasoning at the close of Sect. 5.

But if every sentence says that it itself is true, or implies its own truth, then no sentence is true, or at least no sentence can be shown, using (A) or (A'), to be true. (A) simply says that any such sentence is true if it is true (and whatever else it says or implies is the case). John and Albert are shown to have no adequate theory of truth at all²⁴.

The same problem does not affect Bradwardine's treatment. He is not committed to saying that every sentence says of itself that it is true, only that paradoxical sentences such as *L* (and *C* and *Cq*, if *q* is not true) do so. His theory shows that they are simply not true, and so false, since he accepts Bivalence.

REFERENCES

- ALBERT OF SAXONY. (1974). *Perutilis Logica*. Venice 1522, repr. Olms.
- BURIDAN, J. (2001). *Summulae de Dialectica*. Klima, G. (Tr.). New Haven, CT: Yale University Press.
- CARNAP, R. (1963). "Intellectual autobiography in P. Schilpp". In: *The Philosophy of Rudolf Carnap*. (pp. 3-84). La Salle, IL: Open Court.
- COFFA, J. A. (1991). *The Semantic Tradition from Kant to Carnap: To the Vienna Station*. Cambridge: Cambridge University Press.
- CHURCH, A. (1940). "A formulation of the simple theory of types". In: *Journal of Symbolic Logic*, vol. 5, p. 56-68.
- DAVIDSON, D. (1969). "True to the facts". In: *Journal of Philosophy*, vol. 66, p. 304-323. Reprinted in his *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press, 1984. p. 37-54.
- DODD, J. (2001-2002). "Is truth supervenient on being?" In: *Proceedings of the Aristotelian Society*, vol. 102, p. 69-86.
- HINTIKKA, J. (1975). "A counterexample to Tarski-type truth-definitions as applied to natural languages". In: *Philosophia*, vol. 5, p. 207-212.
- HORWICH, P. (1993). "In the truth domain". In: *Times Literary Supplement* 4711, p. 28.
- _____. (1998). *Truth*. 2d. Ed. Oxford: Clarendon Press.
- INSOLUBILIA. (2009). "Inncyclopedia Britannica". In: Roure, M. L. *La problematique des propositions insolubles au XIIIe siecle et au debut*

²⁴ A fuller treatment of John's, Albert's and Thomas' attempts at a solution was given in (READ, S. "The liar paradox from John Buridan back to Thomas Bradwardine". In: *Vivarium*, 2002. vol. 40, p. 189). I am grateful to John Corcoran and Ignacio Jané, among others at a workshop at the University of Santiago, Spain, for helpful comments on an earlier version of this paper and o Thomas Schmidt and others at the University of Göttingen, Germany, for further helpful comments.

- du XIV^e, suivie de l'édition des traités de W. Shyreswood, W. Burleigh et Th. Bradwardine. *In*: Archives d'Histoire Doctrinale et Littéraire du Moyen Âge, vol. 37, p. 205-326.
- KRIPKE, S. (1975). "Outline of the theory of truth". *In*: Journal of Philosophy, vol. 72, p. 690-716.
- MOODY, E. A. (1953). Truth and Consequence in Medieval Logic. New York: Greenwood Press.
- PRIOR, A. (1962). "Some problems of self-reference in John Buridan". *In*: Proceedings of the British Academy, vol. 48, p. 281-296.
- READ, S. (2002). "The liar paradox from John Buridan back to Thomas Bradwardine". *In*: Vivarium, vol. 40, p. 189-218.
- TARSKI, A. (1936). "The establishment of a scientific semantics" (Tr.). *In*: WOODGER, J. (Ed.). Logic, Semantics, Metamathematics. 2d Ed. Corcoran, J. (Tr.). Indianapolis: Hackett, 1983. p. 401-408.
- _____. (1983). Logic, Semantics, Metamathematics. 2d Ed. Woodger, J. (Ed.). Corcoran, J. (Tr.). Indianapolis: Hackett.
- WITTGENSTEIN, L. (1961). Tractatus Logico-Philosophicus. Pears, D. & McGuinness, B. (Trs.). London: Routledge and Kegan Paul.
- WRIGHT, C. J. G. (1992). Truth and Objectivity. Cambridge, MA: Harvard University Press.
- YAQUB, A. (1993). The Liar Speaks the Truth. Oxford: Oxford University Press.

LOS DEÍCTICOS. UN PROBLEMA PARA LA SEMÁNTICA DE GOTTLÖB FREGE

INDEXICALS. A PROBLEM FOR GOTTLÖB FREGE'S SEMANTIC

ÁNGELA ROCÍO BEJARANO CHAVES
Pontificia Universidad Javeriana, Colombia. annia26@hotmail.com

RECIBIDO EL 12 DE NOVIEMBRE DE 2010 Y APROBADO EL 17 DE DICIEMBRE DE 2010

RESUMEN ABSTRACT

La tesis de este artículo es que tenemos razones suficientes para considerar los deícticos un problema para la propuesta semántica de Gottlob Frege. Dividiremos el texto en dos partes: en la primera, expondremos el programa semántico del lógico alemán por medio de tres tesis estructurantes. En la segunda, introduciremos la cuestión de los deícticos, explorando en qué medida representan un problema para dicho programa.

The thesis of this article is that we have enough reasons to consider indexicals as a problem for Gottlob Frege's semantic. We will divide the text into two parts: in the first one, we will expose the semantic program of the german logician by means of three structuring theses. In the second one, we will introduce the question of the indexicals, exploring how they represent a problem for the above mentioned program.

PALABRAS CLAVE KEY WORDS

deícticos, Frege, Perry, semántica.

Frege, indexicals, Perry, semantic.

Aunque la cuestión de los deícticos aparezca, explícitamente, en los escritos posteriores de Frege, tenemos razones suficientes para considerar la homogeneidad del planteamiento semántico fregeano a lo largo de sus escritos. Por eso, examinaremos la cuestión de los deícticos a la luz de tres tesis semánticas estructurantes y, con ello, expondremos las razones que tenemos para considerar los deícticos un problema para la propuesta semántica de Frege.

I. LA SEMÁNTICA FREGEANA EN TRES TESIS

Las tesis que tendremos en cuenta, y que ofrecen una imagen de conjunto de la propuesta fregeana, son las siguientes: la primera, que la semántica es homogénea; la segunda, que el sentido es distinto a la referencia y a la representación; y la tercera, que el sentido determina la referencia.

1. La semántica es homogénea

La tesis de la homogeneidad de la semántica es que toda expresión con significado significa del mismo modo¹. Aunque en un lenguaje es posible reconocer distintos tipos de expresiones, Frege considera que la diversidad de expresión no es razón suficiente para introducir diversidad en los modos de significar. Por eso, parte de sus mayores esfuerzos teóricos se concentran en explicar cómo la diversidad de expresiones es compatible con un mismo modo de significar. Para empezar, veamos los distintos tipos de expresiones que Frege reconoce.

La semántica fregeana adopta la distinción fundamental entre expresiones incompletas o insaturadas y expresiones completas o saturadas². A las primeras, Frege las llama 'funciones'; a las segundas,

¹ El propósito de Frege es construir un sistema capaz de dar cuenta del pensamiento (FREGE, Gottlob. "Begriffsschrift. A formula language of pure thought modelled on that of arithmetic". En: BEANEY, M. (Ed.). *The Frege Reader*. Oxford: Blackwell Publishers, 1997. pp. Iss.), y aunque éste se nos haga asible por medio de las oraciones del lenguaje, no toda oración expresa pensamientos. Por eso Frege se ocupa de aquellas que sí lo hacen, aquellas relevantes para las inferencias lógicas, las que expresan las verdades de la ciencia: las oraciones afirmativas (FREGE, Gottlob. "El Pensamiento. Una Investigación Lógica". En: VALDÉS, M. (Comp.). *Pensamiento y lenguaje: Problemas en la atribución de actitudes proposicionales*. Ciudad de México: UNAM, 1996. p. 62), y por eso reconoce como expresión, sólo aquellas partes de un enunciado afirmativo relevantes para las inferencias.

² Las expresiones refieren a entidades; las saturadas a entidades saturadas, objetos; y las insaturadas, a entidades insaturadas, funciones. Entonces, así como existen sólo dos tipos de expresiones, existen, para Frege sólo dos tipos de entidades: las funciones y los objetos.

'argumentos'³. Un argumento es una expresión completa; una función es una expresión incompleta, tiene lugares vacíos que pueden llenarse con distintos argumentos^{4,5}. Veamos un ejemplo: si analizamos la expresión 'Homero Simpson es gordo', podemos identificar dos componentes: 'Homero Simpson' y 'es gordo'. El primero, 'Homero Simpson', es una expresión completa, es el nombre propio de algo. El segundo, 'es gordo', es una expresión incompleta, necesita otra expresión que logre completarla, tiene un lugar vacío que puede llenarse con distintos nombres. Podríamos entonces expresarla como 'x es gordo', donde 'x' representaría cualquier nombre que pueda completarla. Así, reconocemos en la expresión compleja original dos tipos de expresiones más simples que la componen: una que completa o satura y una que es completada o saturada. La que satura, 'Homero Simpson', es el argumento; la que es saturada, 'x es gordo', es la función. Ahora bien, si saturamos o completamos la función 'x es gordo' con el argumento 'Homero Simpson' obtenemos la oración inicial 'Homero Simpson es gordo'. La función no saturada logra saturarse, y el resultado es una expresión completa y saturada⁶ que Frege llama 'nombre'⁷. Según esto, un nombre propio como 'Homero Simpson' y una oración como 'Homero Simpson es gordo' son ambas expresiones saturadas o completas, las dos son nombres⁸. La primera sería un nombre simple, la segunda un nombre complejo.

Este análisis deja ver dos características básicas de la homogeneidad semántica. En primer lugar, que todas las expresiones del lenguaje se clasifican según la dicotomía función/argumento. Tal distinción es exhaustiva para las expresiones de un lenguaje. En segundo lugar, que cada expresión del lenguaje pertenece a una y sólo una de estas categorías. Una expresión o es función o es argumento, pero no puede ser función y argumento al mismo tiempo, en este sentido la dicotomía es excluyente⁹. Sin embargo, el punto central de la homogeneidad semántica es que estos dos tipos de expresiones significan de la misma

³ Cfr. FREGE, Gottlob. "Función y Concepto". Moulines, U. (Tr.). *En*: Estudios sobre Semántica, 2a. Ed. Barcelona: Ariel, 1973, p. 23.

⁴ *Ibidem*.

⁵ Una función que pueda completarse con un argumento es una función de primer orden. Con más de un argumento es una relación, y con otra función es una función de segundo orden.

⁶ *Ibid.*, p. 34.

⁷ FREGE, Gottlob. "Sobre sentido y referencia". Moulines, U. (Tr.). *En*: Estudios sobre Semántica, 2a. Ed. Barcelona: Ariel, 1973, p. 51.

⁸ SLUGA, Hans. Gottlob Frege. The arguments of the philosophers. New York: Routledge, 1980. p. 160.

⁹ FRÁPOLLI, M. Una aproximación a la filosofía del lenguaje. Madrid: Síntesis, 1998. p. 50.

manera¹⁰. El proyecto de una semántica homogénea abarca todas las expresiones del lenguaje, sean saturadas o no. En ese orden de ideas, el argumento 'Homero Simpson', la función 'x es gordo', y la oración 'Homero Simpson es gordo', significan, las tres, del mismo modo. La semántica homogénea sostiene que para todas las expresiones del lenguaje el significado funciona de la misma manera. ¿Cuál?

2. El sentido es distinto a la referencia y a la representación

Para Frege, la noción de 'significado' es compleja y no puede reducirse a una sola noción más básica. Por ello propone, para aclararla, dos nociones distintas: 'sentido' y 'referencia'. Una expresión del lenguaje significa en virtud de su sentido y su referencia¹¹. El sentido es "el modo de darse lo designado"¹² o apuntado por la expresión; la referencia es lo referido, lo apuntado¹³. Veamos un ejemplo: las expresiones 'Bruce Dickinson', 'el vocalista de Iron Maiden' y 'el líder de una banda británica', se refieren, todas, a la misma persona. Pero aunque las tres denoten al mismo individuo, no parecen decir lo mismo de él. Parece que las tres mostraran cierta cara o cierta faceta de un mismo individuo. El individuo puede presentarse como 'Bruce Dickinson', como 'el vocalista de Iron Maiden' o como 'el líder de una banda británica'. Pues bien, esos modos de presentarse es lo que Frege llama 'sentidos', y el individuo al que las expresiones denotan es la 'referencia'. En este caso, las tres expresiones en cuestión tendrían la misma referencia, pero distintos sentidos¹⁴.

Del mismo modo en que el sentido del nombre 'Bruce Dickinson' es el modo en el que éste se da, el sentido de la oración 'Bruce Dickinson es inglés' es un pensamiento. El pensamiento, para Frege, es el sentido de

¹⁰ *Ibidem*.

¹¹ BELL, D. Frege's theory of judgement. New York: Oxford University Press, 1979. p. 14.

¹² FREGE, G. Sobre sentido y referencia. Op. cit., p. 50.

¹³ En el caso de los nombres o expresiones saturadas la referencia es un objeto (*Ibid.*, p. 52). De la misma forma en la que el nombre 'Bruce Dickinson' se refiere a Bruce Dickinson, la oración 'Bruce Dickinson es inglés' se refiere a un objeto, y sólo hay dos objetos de los que una oración puede ser nombre: lo verdadero y lo falso (*Ibid.*, p. 60). Entonces, el sentido de una oración es un pensamiento, y la referencia de ésta es un valor de verdad.

¹⁴ Para Frege en un lenguaje ideal a cada nombre le correspondería un sentido y a éste una referencia (*Ibid.*, p. 52). Sin embargo, en el lenguaje natural las relaciones entre los nombres, el sentido y la referencia no tienen tal perfección. A un nombre puede corresponderle una referencia y distintos sentidos, como es el caso mencionado. Por otra parte, en el lenguaje natural, pueden haber nombres con sentido y sin referencia, como es el caso de la expresión 'el príncipe de Argentina', que aunque tenga sentido y podamos entenderla, no tiene referencia.

una oración¹⁵. Pero es necesario aclarar, sin hacer una caracterización exhaustiva de la noción fregeana de 'pensamiento', que ésta no se refiere a una actividad subjetiva o mental de los hablantes. En vez de ello, hace referencia a aquello que puede ser captado por muchos y que no requiere un único portador¹⁶. El pensamiento, para Frege, no es de naturaleza psicológica, ni como proceso ni como contenido. Lo que psicológicamente conocemos como pensamiento, Frege lo llama representación. Por eso el sentido difiere tanto de la referencia como de la representación.

Hasta aquí, tendremos en cuenta dos cosas. Primero, que el sentido es distinto a la referencia. Segundo, que el sentido y, con ello, el pensamiento, no son privados; por eso el sentido es distinto a la representación. Para continuar, expondremos una de las relaciones que Frege establece entre sentido y referencia.

3. El sentido determina la referencia¹⁷

Para Frege, es imposible tener el conocimiento completo de una referencia¹⁸, pero sí es posible que conozcamos varios de los sentidos en los que ella se da. Siendo esos sentidos lo único con lo que contamos, en principio, sólo podemos, a partir de ellos, iluminar la referencia. El hablante conoce sólo un modo de darse de la referencia; así, cuando entiende una expresión del lenguaje, no está conociendo la referencia de dicha expresión, sino un sentido de ésta. Por eso, según Bell¹⁹, 'sentido' es una noción epistemológica, es decir que da cuenta de aquello que conoce un hablante cuando entiende una expresión del lenguaje²⁰.

No obstante, cuando el ser humano se pregunta por la verdad quiere más que sentido. Se pregunta por la referencia. La pregunta por la verdad hace que los hombres vayamos del sentido a la referencia²¹. Así las cosas,

¹⁵ *Ibid.*, pp. 58ss.

¹⁶ *Ibid.*, p. 60.

¹⁷ Aunque Frege no usa la expresión 'determinar', Dummett hace uso de ella para explicar una de las relaciones que se dan entre el sentido y la referencia (DUMMETT, M. *Frege Philosophy of Language*. New York: Harper & Row Publishers, 1973. p. 293, 589). Sin embargo, para Evans la expresión lleva a Dummett a conclusiones erróneas sobre el programa semántico fregeano (EVANS, G. *Ensayos filosóficos*. Ciudad de México: UNAM, 1996. pp. 322ss.). En este artículo nos limitaremos a mencionar la polémica, por lo tanto no entraremos en detalle.

¹⁸ *Ibid.*, p. 52.

¹⁹ BELL, D. *Frege's theory of judgement*. Op. cit., p. 53.

²⁰ DUMMETT, M. *Frege Philosophy of Language*. Op. cit., p. 293.

²¹ FREGE, G. *Sobre sentido y referencia*. Op. cit., p. 60.

teniendo el sentido pero no la referencia, la pregunta por la verdad hace que el sentido sea el camino que lleva a la referencia, es la dirección que conduce a los objetos referidos por cada expresión del lenguaje. Siendo el camino que conduce a la referencia el sentido determina la referencia.

Esa relación de determinación tiene una sola dirección. Va del sentido a la referencia, y no de ésta a aquél. Si la referencia pudiera determinar el sentido, conoceríamos, antes, la referencia, y ésta sería la dirección que nos conduciría a todos los sentidos en los que ésta se da²². Pero eso iría en contra de la propuesta fregeana²³.

II. ¿Y LOS DEÍCTICOS?

Después de exponer las tres tesis de la semántica fregeana, vamos a introducir la cuestión de los deícticos, explorando en qué medida representan un problema para la propuesta del lógico alemán. En primer lugar, intentaremos exponer brevemente la propuesta fregeana en torno a dichos términos; y en segundo lugar, examinaremos los problemas que pueden suscitar para las tesis fregeanas enunciadas.

Los deícticos son expresiones cuyo valor semántico está determinado por rasgos del contexto en el que se emite el término. Pueden ser pronombres personales, como 'yo' o 'tu'; expresiones temporales como 'hoy' y 'ahora'; o locativas como 'aquí' y 'allá', entre otras.

1. Los deícticos en Frege

Para Frege, recordemos, el pensamiento es algo que todos podemos captar, no es privado, y es, también, el sentido de una oración²⁴. Por lo tanto, es lo que conocemos cuando entendemos una oración²⁵. Ahora bien, el pensamiento está determinado por los sentidos de las partes que componen la oración que lo expresa. Si una de esas partes es un término deíctico, es necesario conocer las circunstancias que dicho

²² DUMMETT, M. Frege Philosophy of Language. Op. cit., p. 91.

²³ En primer lugar, porque no podemos conocer directamente una referencia, sino que conocemos los modos en los que ésta se da, los sentidos (FREGE, G. Sobre sentido y referencia. Op. cit., pp. 51ss.). En segundo lugar, porque conocer totalmente una referencia significaría saber inmediatamente si cada sentido dado le pertenece o no, y eso, para Frege, es imposible (*Ibid.*, p. 54). Entonces, el sentido determina la referencia, pero la referencia no determina el sentido.

²⁴ FREGE, G. El Pensamiento. Una Investigación Lógica. Op. cit., pp. 26ss.

²⁵ DUMMETT, M. Frege Philosophy of Language. Op. cit., p. 293.

término indica, para captar correctamente el pensamiento expresado²⁶. Es decir, una oración que contenga deícticos, por sí sola, no expresa completamente un pensamiento, sino que se hace necesario conocer la persona, el lugar o el tiempo que el deíctico expresa, para así poder captar correctamente el pensamiento²⁷. Para Frege, los deícticos juegan un papel fundamental para expresar un pensamiento y captar un pensamiento expresado.

No obstante, Frege afirma que aquello que indica el deíctico no es parte del pensamiento, sino de la expresión del mismo. Entonces, el lugar, el tiempo o la persona que indica el deíctico sería un medio para expresar de forma correcta un pensamiento. Si alguien dice 'hoy se vendieron 1.000 boletas para el partido del Barcelona', el día de hoy, que se expresa por medio del deíctico, no es parte del pensamiento que se está expresando, pero sirve para expresarlo, es parte de la expresión. Si otra persona quisiera expresar el mismo pensamiento, pero un día después de la emisión de la frase mencionada, tendrá que cambiar el deíctico 'hoy' por 'ayer', porque aunque el pensamiento que se busca expresar en ambas oraciones es el mismo, si no se cambia el deíctico, se expresarán dos pensamientos distintos.

Por otra parte, Frege advierte que el deíctico 'yo' da lugar a consideraciones distintas²⁸. Así como 'hoy' y 'ayer', 'yo' indica personas distintas, siempre y cuando sea expresado por distintos hablantes. Pero, a diferencia de los primeros, cuando es expresado por el mismo hablante, indica siempre la misma persona. En el análisis de Frege, antes de que un hablante profiera una oración como 'yo soy estudiante', se ha pensado a sí mismo, y en ese pensarse a sí mismo, la persona se ha presentado a sí misma de una forma única, y a nadie más se le presenta de ese modo. Entonces, el pensamiento que dicha persona tiene, al pensarse a sí misma, sólo puede ser captado por ella y no por otros²⁹. Sin embargo, dicha persona puede usar el deíctico 'yo' si aún desea comunicar algo a otros, porque dicho término es susceptible de ser entendido por ellos.

Hasta aquí, hemos expuesto brevemente la posición fregeana sobre los términos deícticos. Hemos dicho que para él son fundamentales para la captación y la expresión del pensamiento; que se caracterizan por tener

²⁶ FREGE, G. El Pensamiento. Una Investigación Lógica. Op.cit., p. 31.

²⁷ *Ibid.*, pp. 31ss.

²⁸ *Ibid.*, p. 32.

²⁹ *Ibid.*, p. 33.

un significado que varía de preferencia en preferencia; y, que el deíctico 'yo' se diferencia de los deícticos de lugar y tiempo pues, aunque al igual que ellos sirve para expresar y captar pensamientos, su significado no varía para distintas preferencias de un mismo hablante. Para terminar, vamos a examinar por qué los deícticos plantean problemas a la propuesta del lógico alemán.

2. Los deícticos contra Frege

En esta última sección defenderé una tesis débil. Si Frege quiere acomodar los deícticos en su semántica, no puede mantener las tres tesis tal y como están, pero podría hacerlo modificándolas.

La tesis de la homogeneidad semántica es que toda expresión significa del mismo modo. Si un deíctico es una expresión significativa del lenguaje, significaría del mismo modo que todas las otras expresiones: tendría sentido y referencia. Sin embargo, cuando usamos un deíctico parece que, intuitivamente, nos dirigiéramos directamente hacia aquello que está indicando el término en cuestión, más que a un posible sentido en el que se nos presentara. Entonces ¿qué noción de 'sentido' aplicaría para los deícticos? Si una oración debe ser completa para que exprese un pensamiento, y los deícticos adquieren su importancia en la expresión de pensamientos³⁰, podríamos pensar que su papel es complementar la expresión de un pensamiento³¹ para que pueda ser captado correctamente. Pero, ¿puede entenderse ese papel complementador como el sentido del deíctico? Si así fuera, tal papel no daría cuenta de su sentido. Más bien sería una regla que vincula la expresión deíctica con la referencia, pero en tanto regla no variaría como sí lo hace el sentido³². Esto abre la alternativa a una referencia variable y un sentido fijo; pero, si el sentido de un deíctico no variara, y tan sólo lo hiciera su referencia, ¿cómo podría un único sentido determinar infinitas referencias? Mejor aún, ¿admitiría Frege una tesis así?

Estas consideraciones motivan una de dos ideas. O que abandonemos la tesis de la homogeneidad o que la asumamos pensando en los deícticos como un caso límite en que sentido y referencia coinciden, son idénticos.

³⁰ *Ibid.*, p. 31.

³¹ PERRY, J. "Frege Sobre los demostrativos". En: VALDÉS, M. (Comp.) *Pensamiento y lenguaje: Problemas en la atribución de actitudes proposicionales*. Ciudad de México: UNAM, 1996. pp. 56ss.

³² *Ibid.*, p. 57.

En este segundo caso queda en cuestión la segunda tesis según la cual la referencia es distinta al sentido. La primera conclusión es, por tanto: que o se debilita la primera tesis admitiendo los deícticos como excepción, o se rechaza la segunda tesis. Tomemos la primera opción, la más débil, y aceptemos entonces que el sentido, la referencia y la representación son diferentes.

Del sentido, y con ello del pensamiento, se afirmó que son públicos y que pueden ser captados por muchos³³. Pero, cuando Frege explica las nuevas consideraciones que aparecen por el deíctico 'yo', afirma que un hablante que se piensa a sí mismo, capta un pensamiento que ningún otro puede captar³⁴. Es decir, que ese pensamiento tal y como es captado por el que se piensa a sí mismo, es privado. Y entonces, ¿cómo conciliamos ese pensamiento privado, con la tesis del pensamiento como público? Si el pensamiento no es representación, entonces no hay pensamiento privado. Si esto es así, el deíctico 'yo' no tiene cabida en la semántica de Frege. Entonces, la segunda conclusión es que si modificamos la primera tesis para mantener la segunda, tenemos que admitir pensamientos privados, de nuevo en contra del propio Frege.

La tercera tesis de la semántica fregeana es que el sentido determina la referencia. El sentido es primero en el orden epistemológico. Un hablante capta sentidos y luego, con éstos, determina referencias. Pero los deícticos nos llevan directamente a una referencia³⁵. Cuando alguien usa la palabra 'yo', de inmediato indica a la persona que emitió el vocablo, sin que antes se capte un sentido, como insinúa Frege. Parece que lo que conociéramos cuando entendiéramos dicho término no es un modo de darse de un objeto, sino el objeto. Si esto es así, hay camino directo a la referencia sin mediación del sentido, contra la tercera tesis de Frege.

Podría alegarse que un deíctico es un signo que cuenta con innumerables sentidos, dependiendo del tiempo, lugar o persona que indique. Pero un día o una persona no son sentidos³⁶, sino más bien referencias a las que se relacionan múltiples sentidos. De ser así, en el caso de los deícticos, primero estaría la referencia y luego el sentido; pero en la propuesta fregeana no cabe afirmar que una referencia determine un sentido. La tercera tesis semántica de Frege no parece dar lugar a los deícticos.

³³ FREGE, G. Sobre sentido y referencia. Op. cit., p. 60.

³⁴ FREGE, G. El pensamiento. Una Investigación lógica. Op. cit., p. 33.

³⁵ PERRY, J. Frege Sobre los demostrativos. Op.cit., p. 56.

³⁶ *Ibid.*, p. 57.

En suma, si examinamos el significado de los deícticos en el marco de las tres tesis semánticas fregeanas, encontramos dificultades y problemas en cada una de ellas. Si los problemas que hemos planteado son, en verdad, problemas, podríamos proponer o bien modificar alguna de las tesis, o bien eliminarla. Podríamos admitir, por ejemplo, la segunda y tercera tesis, pero eliminar la primera. Es decir, aceptaríamos que el sentido es distinto a la referencia y que la determina, pero que dicha forma de significar no sería la misma para todas las expresiones del lenguaje. Los deícticos significarían de otros modos.

Por otra parte, podríamos aceptar la homogeneidad semántica, y la distinción entre el sentido y la referencia, pero tendríamos que modificar la tercera tesis, afirmando que existe bidireccionalidad en la determinación. El sentido determinaría a la referencia, y la referencia determinaría al sentido. En este caso, afirmar que el deíctico nos conduce directamente al objeto referido, ya no sería un problema.

Entonces, podríamos eliminar la primera tesis o modificar la tercera. En cualquiera de los dos casos se resolvería la cuestión del significado de los deícticos dentro de la propuesta fregeana. Sin embargo, en cualquiera de estos casos tendríamos que resolver el problema que surge a partir de la segunda tesis, a saber, el pensamiento privado que parece tener una persona cuando se piensa a sí misma. Si queremos mantener la propuesta fregeana del pensamiento público, nos vemos obligados a negar, pese a la propuesta del mismo Frege, que existe un pensamiento que sólo puede ser captado por una persona. Y esa, es una tarea más difícil.

En consecuencia, habría que decir que, si vamos a contemplar los deícticos en una semántica fregeana, las tres tesis no pueden sostenerse al tiempo. En este sentido, los deícticos son un problema para la semántica de Frege o, si se quiere, la semántica de Frege es un problema para los deícticos.

REFERENCIAS

- BELL, D. (1979). *Frege's theory of judgement*. New York: Oxford University Press.
- DUMMETT, M. (1973). *Frege Philosophy of Language*. New York: Harper & Row Publishers.
- EVANS, G. (1996). *Ensayos filosóficos*. Ciudad de México: UNAM.
- FRÁPOLLI, M. (1998). *Una aproximación a la filosofía del lenguaje*. Madrid: Síntesis.
- FREGE, Gottlob. (1997). "Begriffsschrift. A formula language of pure thought modelled on that of arithmetic". En: BEANEY, M. (Ed.). *The Frege Reader*. Oxford: Oxford Blackwell Publishers.
- _____. (1996). "El Pensamiento. Una Investigación Lógica". En: VALDÉS, M. (Comp.). *Pensamiento y lenguaje: Problemas en la atribución de actitudes proposicionales*. Ciudad de México: UNAM.
- _____. (1973). "Función y Concepto". Moulines, U. (Tr.). En: *Estudios sobre Semántica*. 2a. Ed. Barcelona: Ariel.
- _____. (1973). "Sobre sentido y referencia". Moulines, U. (Tr.). En: *Estudios sobre Semántica*. 2a. Ed. Barcelona: Ariel.
- PERRY, J. (1996). "Frege Sobre los demostrativos". En: VALDÉS, M. (Comp.) *Pensamiento y lenguaje: Problemas en la atribución de actitudes proposicionales*. Ciudad de México: UNAM.
- SLUGA, Hans. (1980). *Gottlob Frege. The arguments of the philosophers*. New York: Routledge.

JOHN RAWLS: A LIBERDADE DOS ANTIGOS E A LIBERDADE DOS MODERNOS – PRESSUPOSTOS DA JUSTIFICAÇÃO COERENTISTA

*JOHN RAWLS: THE LIBERTY OF ANCIENTS AND THE LIBERTY
OF MODERNS—PRESUPPOSITIONS OF COHERENTIST JUSTIFICATION*

ELNORA GONDIM

Universidade Federal do Piauí, Brasil. elnoragondim@yahoo.com.br

RECIBIDO EL 14 DE AGOSTO DE 2010 Y APROBADO EL 30 DE SEPTIEMBRE DE 2010

RESUMO ABSTRACT

A ênfase na liberdade individual e na igualdade de todos os cidadãos vista nos dois princípios de justiça rawlsianos faz com que a justiça como equidade efetue uma articulação entre a liberdade individual e a coletiva explicitamente vista nas duas faculdades morais contidas na concepção de pessoa da teoria rawlsiana, isto é, na idéia de racionalidade e na concepção de razoabilidade. Portanto, desta maneira, Rawls elabora uma articulação entre o público e o privado inserindo, assim, na justiça como equidade tanto a tradição antiga de liberdade quanto a moderna. Na justiça como equidade, a utilização do método do equilíbrio reflexivo, a autonomia, a liberdade, o conceito normativo de pessoa, o procedimentalismo puro, a ordem lexicográfica atribuída ao primeiro princípio, à concepção política de justiça e a base pública de justificação, isto tudo mostra que Rawls considera tanto a liberdade dos antigos quanto a dos modernos e, com base neste argumento, constata-se um forte suporte conceitual que aponta para o modelo rawlsiano de justificação: o coerentista.

The emphasis in the individual freedom and the equality of all citizens as it appears in the two Rawlsian principles of justice causes that justice as fairness be part of an articulation between the individual freedom and the collective one as explicitly seen in the two moral faculties contained in the conception of person of the Rawlsian theory, that is, in the idea of rationality and in the conception of reasonableness. Therefore, Rawls elaborates an articulation between the public and the private inserting justice as fairness in such a way that he is able to integrate the old tradition of freedom as well as the modern one. In justice as fairness, the use of the method of reflective equilibrium, autonomy, freedom, the normative concept of person, pure proceduralism, the lexicographical order attributed to the first principle, the political conception of justice and the public base of justification, show that Rawls considers the freedom of the ancients in the view of the freedom of the moderns and, on the basis of this argument, provides strong conceptual support for his model of justification: a coherentist one.

PALAVRAS CHAVE KEY WORDS

liberdade, antigos, modernos, coerentismo.

freedom, ancients, moderns, coherentism.

INTRODUÇÃO

Edmund Gettier, em 1963, publicou seu famoso artigo *É o Conhecimento Crença Verdadeira Justificada?* É consenso que este pequeno texto modificou os rumos tomados pela epistemologia, porquanto analisou a afirmação em relação à verdade justificada do conhecimento proposicional, aquele que é considerado como o saber acerca dos fatos. Gettier constatou que havia uma falha na definição tradicional de conhecimento como crença verdadeira justificada: possuir uma crença verdadeira justificada pode não ser conhecimento; para ele, é possível não possuir qualquer conhecimento mesmo que se tenha uma crença verdadeira justificada. Portanto, após o artigo *É o Conhecimento Crença Verdadeira Justificada?* tem-se a necessidade do surgimento de teorias que exploram as causas ou produções das crenças tanto ao nível de epistemologia quanto ao nível da moral. Depois do texto de Gettier é inconcebível que não se coloque como problema à questão de justificabilidade das crenças. Em decorrência disto foi desencadeada a necessidade da justificativa epistêmica, também, no âmbito da moralidade. As justificações mais comuns referentes à teoria moral são: o fundacionismo e o coerentismo. Têm-se como tipos mais freqüentes de fundacionismo moral: o intuicionismo e o utilitarismo. Contraopondo-se a estas correntes, Rawls utilizou, na sua teoria da justiça, do procedimento coerentista para justificar princípios e juízos morais.

Saber a importância de se mostrar o tipo de justificabilidade para a ação moral é algo já aqui enfatizado: o escrito de Gettier exigiu tal ênfase. Mas, por que a teoria rawlsiana tem uma importância ao ponto de se discutir qual o tipo de justificação que nela é utilizada? Por conseguinte, por que se deve saber qual a forma de Rawls lidar com a ação moral?

Em se tratando de Rawls, com ele ocorreu algo semelhante a Gettier. Uma Teoria da Justiça, obra publicada em 1971, inova o contexto da teoria moral e política contemporânea fazendo emergir, a partir dela, muitas polêmicas e críticas tais como, por exemplo, a dos comunitaristas. Estas podem ser sintetizadas em cinco teses: 1) uma concepção abstrata de pessoa que é consequência da posição original sob o véu da ignorância; 2) princípios universais (deontológicos) com a pretensão de aplicação em todas as sociedades, criando uma supremacia dos direitos individuais em relação aos direitos coletivos; 3) não possui uma teoria da sociedade em função de seu contratualismo, trazendo como consequência uma atomização do social, em que a pessoa é considerada enquanto

átomo isolado; 4) utiliza a idéia de um Estado neutro em relação aos valores morais, garantindo apenas a autonomia privada (liberdade dos modernos) e não a autonomia pública (liberdade dos antigos), estando circunscrita a um subjetivismo ético liberal; 5) é uma teoria deontológica e procedimental, estabelecendo uma prioridade absoluta do justo em relação ao bem¹; além destas críticas, apontam, também, a teoria rawlsiana como um intuicionismo (fundacionismo) disfarçado². Portanto, em virtude das polêmicas e críticas, John Rawls passa a ser considerado, após a publicação do texto de 1971, um reabilitador das questões morais e políticas. Para se compreender em que medida Rawls reabilita tais questões é necessário constatar a relevância que a concepção de liberdade assumiu no curso da suas obras. Tal concepção permeia todo o conjunto das obras rawlsianas corroborando com: (i) as respostas dadas por Rawls aos seus inúmeros críticos; (ii) a tendência rawlsiana à justificação do tipo coerentista para a ação moral.

1. LIBERDADE DOS ANTIGOS E LIBERDADE DOS MODERNOS

A distinção entre aquilo que é designado como liberdade dos antigos e dos modernos remonta a conferência escrita por Benjamin Constant em 1819. Na teoria rawlsiana isto toma a forma de um sólido suporte conceitual. Embora Rawls afirme que a oposição efetuada por Constant entre a liberdade dos antigos em relação aos modernos seja algo vago e inexato, a justiça como equidade considera tanto a liberdade negativa quanto à positiva³. Tendo em vista tal aspecto, constata-se a presença contínua, concomitante, nas obras rawlsianas das noções de liberdade antiga e moderna, isto é, da menção: (i) à igualdade; (ii) à liberdade juntamente com o procedimentalismo puro e com a elaboração dos princípios de justiça como algo histórico e analítico; (iii) à crença na liberdade individual enfatizada no primeiro princípio de justiça; (iv) à crença na igualdade de todos os cidadãos em uma sociedade enfatizada no segundo princípio de justiça.

¹ SILVEIRA, Denis Coitinho. “Teoria da justiça de John Rawls: entre o liberalismo e o comunitarismo”. In: *Trans/Form/Ação*, 2007. vol. 30, no. 1, p. 169.

² E. BONELLA, Alcino. *Justiça como imparcialidade e contratualismo*. Campinas, IFCH- Unicamp. (Tese de Doutorado), 2000.

³ Liberdade positiva ou dos antigos é a participação coletiva no exercício da soberania. Liberdade negativa ou dos modernos é a liberdade privada ou o exercício pelo indivíduo do seu direito natural de gerir sua vida como bem entende.

Quanto aos princípios de justiça, estes são: (i) denominado princípio da igual liberdade; assegura certas liberdades básicas iguais a todos os cidadãos. (ii) denominado princípio da igualdade eqüitativa de oportunidades e da diferença; requer o Estado como regulador ao nível de distribuição de riquezas levando em conta e priorizando os menos favorecidos. Portanto, a ênfase na liberdade individual e na igualdade de todos os cidadãos vista nos dois princípios de justiça rawlsianos faz com que a justiça como eqüidade efetue uma articulação entre a liberdade individual e a coletiva explicitamente vista nas duas faculdades morais contidas na concepção de pessoa da teoria rawlsiana, isto é, na idéia de racionalidade (aquela que é relativa ao bem, tem uma forma de privado) e na concepção de razoabilidade (aquela que é relativa ao justo, tem uma forma de público). Assim sendo, através do razoável os indivíduos são iguais no mundo público dos outros e podem propor, aceitar e dispor termos eqüitativos de cooperação entre eles, em contrapartida em virtude do racional nenhum indivíduo é levado a desistir de suas concepções do bem. Portanto, desta maneira Rawls elabora uma articulação entre o público e o privado inserindo, assim, na justiça como eqüidade tanto a tradição antiga de liberdade quanto a moderna.

1.1 A liberdade dos antigos

Com a ênfase dada pela justiça como eqüidade à liberdade antiga, cumpre aqui se tentar sintetizá-la. Na *Pólis* o homem não era visto como indivíduo, mas como membro da comunidade; ele tinha direitos políticos e isto constituía a vontade coletiva, onde todas as virtudes e aptidões eram realizadas no e para o Estado. Aqui é interessante lembrar que o substrato da democracia grega era a noção de comunidade. Com isso, havia uma negação para o indivíduo de toda uma idéia de autonomia subjetiva e de direitos individuais perante a comunidade. Assim, na liberdade dos antigos, ao cidadão era dado participar da elaboração e da direção do poder estatal tendo ele a mesma igualdade no exercício da vida pública sem reflexos imediatos no plano das garantias da vida privada. Desta forma, essa liberdade acontecia em sua plenitude quando o cidadão afirmava a sua vontade dentro dos limites da *Pólis*.

Aqui cumpre salientar que a liberdade dos antigos, na visão de Constant, está relacionada à perspectiva teórica de Rousseau e, neste caso, designa as liberdades políticas iguais e os valores da vida pública. Para a filosofia rousseauniana, o Estado é aquele que cria, a partir do homem natural, um ser novo a quem deverá ensinar tudo; o homem, em sua origem,

é apenas um animal. Sendo assim, é preciso fazer dele um cidadão e, com isto, o Estado, então, encontra-se em um nível de superioridade em relação ao indivíduo e o papel dele é educar e emancipar tal indivíduo fazendo-o livre. Desta maneira, para Rousseau, o homem nem sempre vê o seu próprio bem e a prioridade dada aqui é para a igualdade das liberdades políticas e para os valores da vida pública, subordinando, assim, as liberdades civis.

De Rousseau, Rawls tem influências quanto ao tema da liberdade e da autonomia; defendendo uma concepção de justiça como imparcialidade e não um Estado do bem-estar que parte de paradigmas fundacionais pertinentes ao bem. Neste contexto, é sob a influência de Rousseau, que Rawls utiliza a idéia de intenção manifesta, isto é, aquela que pode ser definida como: os cidadãos se dispõem de boa vontade a fazer sua parte nos arranjos desde que tenham uma garantia suficiente de que os outros também farão a sua. Rawls afirma que Rousseau tratou a desigualdade de maneira significativa. A solução que é vista na filosofia rousseauiana para isto é análoga na justiça como equidade, porém com modificações, pois: “o status fundamental na sociedade política é a cidadania igual para todos, um status que todos tem como pessoas livres e iguais (...) É do ponto de vista de cidadãos iguais que a justificação de outras desigualdades deve ser entendida”⁴.

Nestas circunstâncias, Rawls afirma que todos os cidadãos são iguais e isto favorece o relacionamento igualitário entre os mesmos de forma que todos devem ser favorecidos por suas políticas públicas. Isto é visto na teoria rawlsiana no aspecto em que esta afirma que a liberdade é agir de acordo com a lei que estabelecemos para nós mesmos. Assim sendo, Rawls afirma: “Não importa qual seja a posição de uma pessoa no tempo, cada uma é forçada a escolher por todos”⁵. Sendo desta forma, o pensamento rawlsiano segue o de Rousseau, mais precisamente quando este afirma:

Deve-se compreender, neste sentido, que, menos do que o número de votos, aquilo que generaliza a vontade é o interesse comum que os une, pois nessa instituição cada um necessariamente se submete às condições que impõe aos outros: admirável acordo entre o interesse e a justiça, que dá às deliberações comuns um caráter de equidade que vimos

⁴ *Ibid.*, Justiça como equidade; uma reformulação, p. 186.

⁵ *Ibid.*, TJ, p. 151.

desaparecer na discussão de qualquer negócio particular, pela falta de interesse comum que uma e identifique a regra do juiz à parte (...) o pacto social estabelece entre os cidadãos uma tal igualdade que eles se comprometem todos nas mesmas condições e devem gozar dos mesmos direitos⁶.

Desta forma, o neocontratualismo de Rawls “reside sobretudo em seu (...) liberalismo constitucional (reapropriação dos ideais (...)), de igualdade e vontade de Jean Jacques Rousseau).”⁷ Neste sentido, o liberalismo político de Rawls tem como tema central o igualitarismo de inspiração rousseauiana, decorrendo dele o princípio de equidade, da igual liberdade, articulando, desta forma, Vontade Geral e igualdade política, promovendo a liberdade igual para todos e a equidade de oportunidades, tendo uma concepção de justiça como imparcialidade e não defendendo um Estado do bem-estar nem um modelo comunitário deste, embora: “(...) a concepção da prioridade do direito de Rawls é mais exigente que a abordagem (...) de Rousseau, no sentido de que ele defendeu não apenas a igualdade, mas igualdade máxima de direitos, assumindo assim que o direito tem uma métrica (...)”⁸. E para isto, Rawls parte de teoria do contrato social como a de Rousseau para chegar a um modo de representação relacionada com a idéia de que os princípios de justiça são objetos de um acordo racional deliberados através de um procedimento neutro e imparcial tal qual o recurso rousseauiano do caráter impessoal da lei e “(...) Daí para o véu de ignorância de Rawls não há mais do que um passo. A desigualdade é inevitável, só se trata de regulá-la, de transformá-la num mecanismo impessoal que não marque desde o início o destino de cada sujeito(...)”⁹.

1.2 A liberdade dos modernos

Quanto à liberdade dos modernos: aquela que resguarda o indivíduo do poder estatal, ela é ligada ao direito; este tem como objetivo intervir no Estado a favor dos interesses individuais. Assim sendo, na liberdade moderna há uma valorização consciente e deliberada do indivíduo perante a sociedade estatal. Desta forma, os modernos se perguntavam sobre aquilo que consideravam prescrições impositivas da justa razão

⁶ ROUSSEAU. “Do contrato social”. In: Coleção Os Pensadores. São Paulo: Abril Cultural, 1978. p. 30.

⁷ OLIVEIRA, Nythamar. Rawls. Rio de Janeiro: Zahar, 2003. p. 17.

⁸ O’NEILL, Onora. Em direção à justiça e à virtude: uma exposição construtiva do raciocínio prático. São Leopoldo: Unisinos, 2006. p. 192.

⁹ BORON, Atilio. A. Filosofia política moderna: de Hobbes a Marx. [on line] [http:// bibliotecavirtual.clacso.org.ar/ar/libros/secret/filopolmpt/filopolmpt.html](http://bibliotecavirtual.clacso.org.ar/ar/libros/secret/filopolmpt/filopolmpt.html), p.93.

sobre seus direitos, seus deveres e às suas obrigações às quais suas prescrições davam origem. Somente depois disso, é que eles voltavam sua atenção aos bens que estas prescrições permitiam buscar e apreciar.

A liberdade dos modernos, segundo Constant, é oriunda de Locke e tem como significado aquela de pensamento, de consciência; certos direitos básicos de pessoa, de propriedade e o primado da lei. Para Locke, o estado de natureza é uma situação de liberdade e igualdade; sendo a liberdade entendida como: “(...) a liberdade consiste em não se estar sujeito à restrição e à violência por parte de outras pessoas (...)”¹⁰. Quanto à igualdade, segundo Locke, esta é jurídica e : “(...) consiste, para cada homem, em ser igualmente o senhor de sua liberdade natural, sem depender da vontade nem da autoridade de outro homem (...)”¹¹. Contudo, no estado de natureza, como o homem faz a sua própria justiça, isto pode gerar conflitos e um estado de guerra. Logo, no estado de natureza há a inconveniência da ausência de um juiz imparcial. Desta forma, falta uma instituição capaz de julgar as controvérsias que aparecem. Neste sentido, o estado civil aparece para sanar as inconveniências do estado de natureza e o Estado é uma instituição com o objetivo de possibilitar à convivência natural entre os homens, fazendo serem respeitadas as leis naturais.

Em Locke, antes do contrato, os homens no estado de natureza já são dotados de todas as suas faculdades. Neste sentido, o papel do Estado é limitado; a ele só cabe garantir a segurança individual sem intrometer-se em sua esfera privada. No entanto, o homem deve obedecer ao governo como, também, ao primeiro princípio da lei natural que é aquele da liberdade, onde as instituições coercitivas se justificam quando as mesmas promovem a liberdade. Assim sendo, a proteção dos direitos fundamentais como o direito à vida, à liberdade e à propriedade não é renunciado; o que o indivíduo renuncia é o direito de fazer justiça por si mesmo. Em consequência disto, para Locke, a liberdade é relacionada com os direitos naturais e para ela ser efetivamente realizada é preciso um consenso onde todos consentam em se submeter às determinadas leis. Assim, a liberdade natural, para Locke, deveria vir antes do que à ordem, isto é, a ordem, nesse caso, seria um meio para fazer prevalecer à liberdade e quando a ordem é opressora, a liberdade tem prioridade. Para que ocorra a paz, os homens têm que ter liberdade religiosa, onde

¹⁰ LOCKE, John. Segundo Tratado Sobre o Governo Civil: e outros escritos. In: Clássicos do Pensamento Político. Petrópolis: Vozes, 1994. p. 115.

¹¹ *Ibid.*, p. 114.

isto envolve a questão da tolerância. Quanto a esta, Locke argumenta que tem de haver uma separação entre a sociedade política e a igreja, embora a origem de ambas seja a mesma

Quanto a Rawls, este retoma de Locke a idéia da desobediência civil, da liberdade e da tolerância (provavelmente, aquilo que possibilita a racionalidade, estrutura a razoabilidade, unindo-as, tornando-as inseparáveis enquanto idéias complementares, fazendo com que o método coerentista rawlsiano, equilíbrio reflexivo *wide*, não priorize os fundamentos de nenhuma doutrina compreensiva). Porém, quando Rawls afirma que a posição original generaliza a idéia do contrato social e o faz constituindo em objeto do acordo os princípios de justiça para a estrutura básica, isto ele não o faz como Locke, isto é, para uma determinada forma de governo. Para Rawls, a filosofia de Locke submete indevidamente as relações sociais entre as pessoas morais às contingências históricas e sociais estranhas à sua liberdade e à sua igualdade e isto a enfraquece.

Contrariamente a Locke, Rawls elabora uma teoria da justiça como equidade que tem a estrutura básica como objeto primeiro de justiça, desenvolvendo uma idéia de acordo particular dado o caráter e o papel único da estrutura básica da sociedade. Neste sentido, Rawls, em dissonância com Locke, diz que “a posição original generaliza a idéia familiar do contrato social. E o faz constituindo em objeto do acordo com os primeiros princípios da justiça para a estrutura básica e não para uma determinada forma, como em Locke¹². Rawls é, também, contrário à filosofia de Locke porque este, em seu contrato hipotético de caráter histórico, sacrifica a liberdade do ser humano uma vez que, como o homem é afetado diretamente pelas contingências e pelos acidentes, nada leva a garantir para ele a justiça do contexto social, pois, em consequência do pacto social, os membros da sociedade não obtêm todos os direitos políticos iguais, porque estes são diretamente relativos à propriedade. Com isto, a doutrina de Locke submete as relações sociais a estranhamentos em relação à liberdade e à igualdade.

No entanto, Locke prioriza a liberdade de pensamento e de consciência, a propriedade, a associação e a determinados direitos básicos do indivíduo. Deste modo, Rawls, assim como Locke: (i) em relação à teoria da desobediência civil como forma de liberdade, enfatiza a função do

¹² RAWLS, John. Justiça como equidade: uma reformulação. São Paulo: Martins Fontes, 2003. p.23.

contrato e da justiça procedural pura; (ii) quanto à questão da tolerância, vê no contrato um acordo entre pessoas razoáveis e racionais que, embora divergentes em muitos assuntos, podem conviver pacificamente. Nesta perspectiva, em consonância com a filosofia de Locke, Rawls enfatiza os direitos individuais e, para isto, reabilita o conceito de justiça que é inerente ao contratualismo

CONSIDERAÇÕES FINAIS:

Pode-se afirmar que dentro das circunstâncias acima mencionadas, para Rawls:

é o que existe entre a tradição de Locke, que dá mais importância ao que Benjamin Constant chama de “liberdade dos modernos”, isto é, liberdade de pensamento e consciência, certos direitos básicos da pessoa e da propriedade, e a tradição de Rousseau, que põe a ênfase na “liberdade dos antigos”, ou seja, a igualdade das liberdades políticas e os valores da vida pública. Esse contraste é vago e historicamente inexato, mas pode servir para fixar as idéias¹³.

Assim, a concepção de liberdade moderna comparada à antiga tendo como referência, respectivamente, Locke e Rousseau, traz conflitos. Estas concepções revelam uma tensão que decorre dos interesses sociais, econômicos, onde o funcionamento das instituições significa: (i) diferenças entre teorias políticas econômicas e sociais gerais; (ii) concepções diferentes sobre as consequências de políticas públicas. Desta maneira, o conflito é em virtude de como as diferentes doutrinas filosóficas e morais entendem as exigências entre a liberdade e a igualdade, a prioridade entre elas, a sua ordenação e como se deve justificar uma maneira de ordenar tais exigências. Pode-se, então, sintetizar o porquê disto tudo ocorrer; isto ocorre pelo fato do antagonismo constatado nas duas concepções em relação à igualdade e à liberdade; há um debate e não há um acordo público sobre como as instituições básicas devem ser organizadas para melhor se adequarem à liberdade e à igualdade da cidadania democrática.

¹³ *Ibid.*, p.207.

Para Rawls, seria pouco satisfatória somente a acomodação das concepções de liberdade antiga e moderna. Sendo assim, a teoria da justiça como equidade tenta arbitrar entre essas duas concepções propondo dois princípios de justiça para serem guias na efetivação dos valores da liberdade e da igualdade, definindo um argumento que mostre o motivo desses princípios aparecem mais apropriados do que os outros para a natureza dos cidadãos de uma democracia. Rawls, para isto, presumirá que a liberdade pode ser explicada mediante aos seguintes aspectos: agentes que são livres às restrições que apontam para a liberdade deles e aquilo que eles estão livres para fazer ou não. Somente isto é relevante para explicar o que é liberdade. Para isto, a filosofia rawlsiana considera, também, a cultura política pública. Nessa perspectiva, para uma concepção política da justiça ser aceita, esta deve estar de acordo com convicções bem ponderadas depois de uma reflexão ponderada.

Com base no exposto, Rawls afirma que a justiça como equidade considera que a liberdade dos antigos tem menos valor intrínseco que a liberdade dos modernos. Então, em uma sociedade moderna, a política não é o centro de tudo e as liberdades políticas são consideradas básicas sendo meios institucionais essenciais para proteger outras liberdades. Rawls, então, assegura que vai tentar ignorar as discussões sobre o significado da liberdade como, também, dos conflitos entre os proponentes da liberdade negativa e os da positiva, pois, para ele, a liberdade de pensamento e a liberdade de consciência, a liberdade individual e as liberdades civis não deveriam ser sacrificadas em nome da liberdade política. Desta forma, embora seja de real importância o engajamento dos cidadãos na vida política, isto deve ser feito mediante uma escolha deles próprios. Porém, as reivindicações dos indivíduos não podem ter prioridade em relação aos princípios de justiça nem da liberdade que eles garantem.

Com isto, o que Rawls afirma é uma conjugação entre liberdade e igualdade, pois a justiça como equidade não aceita nenhum acordo que viole as liberdades básicas, assegurando, então, que há um direito natural à liberdade. Em Rawls, há uma igualdade democrática na diferença e o que deve ser igual é a realização de um projeto de vida racional considerando, para isto, a liberdade. Neste sentido, os princípios de justiça seguem uma ordem lexicográfica, onde as liberdades básicas têm prioridades sobre o princípio da diferença. Assim sendo, Rawls protege a autonomia individual e a defesa das liberdades deve ser à base de toda justiça.

Em consonância com o acima exposto, Rawls não admite que entre a concepção de liberdade dos antigos e dos modernos exista uma oposição fundamental, pois o que se deve levar em consideração é o grau de engajamento que os cidadãos devem ter na política para garantir suas liberdades básicas e qual a melhor maneira para consegui-las. Para tanto, a teoria rawlsiana leva em consideração tanto a natureza social do cidadão como, também, a sua autonomia. Então, o que ele propõe é que os cidadãos compartilhem de uma cidadania igual, que a liberdade igual seja pública e consensualmente estabelecida através de julgamentos bem ponderados tendo como método o equilíbrio reflexivo para que, desta maneira, ocorra uma limitação na discordância pública tendo como base a tolerância.

Desta forma, para a teoria rawlsiana a concepção política deve combinar idéias e princípios bem conhecidos, mas conectados de maneira nova, levando em consideração que a sociedade é um sistema de cooperação eqüitativa entre pessoas livres, iguais e garantindo, assim, um consenso, onde o objetivo da justiça como eqüidade: “(...) não é nem metafísico nem epistemológico, mas prático (...) é um acordo político informado e totalmente voluntário entre cidadãos que são considerados como pessoas livres e iguais (...)”¹⁴. Assim, para Rawls, tanto a liberdade dos antigos como a dos modernos deve ser levada em consideração, não como uma forma de, simplesmente, sintetizá-las, uni-las, mas, acima de tudo, como uma maneira de fornecer uma base aceitável para as instituições democráticas. Neste sentido, ele afirma que: “As liberdades políticas iguais e a liberdade de pensamento devem assegurar a aplicação livre e bem-informada dos princípios de justiça, por meio do exercício pleno e efetivo do senso de justiça dos cidadãos, à estrutura básica da sociedade”¹⁵.

Aqui convém ressaltar que, para Rawls, as concepções de pessoa e sociedade são levadas em consideração especificando os direitos e deveres básicos que devem ser garantidos pelas instituições políticas e sociais referentes à estrutura básica da sociedade; tudo isto efetivado mediante um processo de construtivismo político que é gerado pelo dispositivo da posição original, véu de ignorância e, sobretudo, pelo método do equilíbrio reflexivo. Neste sentido, é conveniente lembrar que a concepção de pessoa é normativa como, também, política e, por ser assim, ela tem como caracteres constitutivos os dois poderes morais.

¹⁴ RAWLS, John. *Justiça e Democracia*. São Paulo: Martins Fontes, 1998. p. 210.

¹⁵ RAWLS, John. *O liberalismo político*. São Paulo: Ática, 2000. p. 391.

Desta forma, as pessoas são livres e autônomas, não se pautando em nenhum princípio previamente dado e os princípios de justiça são formulados a partir de um levantamento feito em relação a regimes bem-sucedidos, partindo de um exame para se saber quais são os direitos fundamentais. Esta ênfase na concepção de pessoa baseada em seres que examinam e constroem os princípios de justiça, faz com que Rawls não participe das disputas em relação aos valores políticos que caracterizam as controvérsias entre a liberdade dos antigos e dos modernos.

Contudo, embora Rawls não participe das controvérsias em torno dos dois modelos de liberdade, isto não significa que ele não as leve em consideração: quanto à liberdade negativa, a teoria rawlsiana a contempla quando afirma que as liberdades não são absolutas, que elas devem ser avaliadas como um todo, com um princípio único, onde, através deste, pode-se ponderar se há uma melhoria no total das liberdades; aqui cumpre ressaltar que uma liberdade só é restrita em função de outra. Com isto, há uma ausência de restrições externas, uma ausência de impedimentos e o Estado não deve se interpor na esfera da liberdade.

Quanto à liberdade positiva, Rawls, também, a considera. Embora ele afirme que a liberdade não é absoluta, ele parte do pressuposto da necessidade de uma ordem léxica¹⁶ não só quanto aos princípios de justiça, mas, também, no interior do primeiro princípio. Assim, o argumento da liberdade negativa é revertido àquele da liberdade positiva onde as liberdades políticas têm prioridade e são incluídas entre as liberdades básicas. Desta forma, Rawls trata dos direitos básicos, das liberdades essenciais e a justiça como equidade é caracterizada como uma concepção política que tem como objetivo administrar conflitos entre liberdades fundamentais, levando em consideração a autonomia das pessoas como seres livres, iguais que podem rever, construir e avaliar os seus princípios de justiça. Com base nisto, Rawls afirma uma prioridade do razoável em relação ao racional; os cidadãos agindo assim exprimem a sua autonomia completa e ele afirma que:

a meta inicial da teoria da justiça como equidade (...) é mostrar que os dois princípios de justiça permitem compreender melhor as reivindicações da liberdade e da igualdade numa sociedade democrática do que o fazem os

¹⁶ A ordenação lexical confere a todos os constituintes um peso específico sem os tornar mutuamente substituíveis.

primeiros princípios associados às doutrinas fundamentais do utilitarismo, ao do perfeccionismo e do intuicionismo¹⁷.

Portanto, a característica de uma justiça procedimental pura é a especificação do que é justo pelo resultado do procedimento e, sendo assim, tanto o resultado quanto o procedimento formam um sistema consistente, justo e coerente. Então, segundo Rawls:

a justiça como equidade almeja pôr de lado antigas controvérsias religiosas e filosóficas e não se apóia em qualquer visão abrangente específica. Faz uso de uma idéia diferente, a de justificação pública, e procura moderar conflitos políticos irreconciliáveis e determinar as condições para uma cooperação equitativa entre cidadãos. Para realizar esse objetivo tentamos elaborar, a partir de idéias fundamentais implícitas na cultura política, uma base pública de justificação que todos os cidadãos, considerados razoáveis e racionais, possam endossar a partir de doutrinas abrangentes. Caso isto se concretize, temos um consenso sobreposto de doutrinas razoáveis¹⁸.

Logo, na justiça como equidade, a utilização do método do equilíbrio reflexivo, a autonomia, a liberdade, o conceito normativo de pessoa, o procedimentalismo puro, a ordem lexicográfica atribuída ao primeiro princípio, à concepção política de justiça e a base pública de justificação, isto tudo mostra que Rawls considera tanto a liberdade dos antigos quanto à dos modernos e, com base neste argumento, constata-se um forte suporte conceitual que aponta para o modelo rawlsiano de justificação: o coerentista.

Neste sentido, vê-se, corroborado com o acima exposto, que a teoria rawlsiana tem uma justificação coerentista uma vez que, além de não se ater á crenças básicas quanto à elaboração dos princípios de justiça, pois os mesmos são construídos em posição original através do equilíbrio reflexivo, o recurso ao procedimento puro não tem um pressuposto *a priori* e as crenças contidas nele mutuamente se apóiam no sentido de que os critérios e os procedimentos são inseparáveis; eles formam um sistema justo e coerente, onde a decorrência justa que um traz é em consequência do aspecto de justiça do outro.

¹⁷ RAWLS. Justiça e democracia. Op. cit., p. 146.

¹⁸ RAWLS. Justiça como equidade: uma reformulação. Op. cit., p. 40.

Assim, o coerentismo rawlsiano está relacionado à problemática da liberdade negativa e positiva: as partes, no momento da elegibilidade dos princípios de justiça, realçam as suas características da racionalidade (liberdade dos modernos; exemplo disto: cada cidadão pode determinar qual doutrina filosófica, religiosa ou metafísica considera verdadeira) e da razoabilidade (liberdade dos antigos, exemplo: ênfase ao político, ao justo). Estes aspectos indicam que, no construtivismo rawlsiano, o consenso sobreposto, realizado através do equilíbrio reflexivo e de convicções ponderadas, resgata aspectos relativos aos dois modelos de liberdade, gerando uma síntese, caracterizando, através deste procedimento, um recurso coerentista para a justificação rawlsiana: o do caráter holístico da mesma, ou seja, ao afirmar que a deliberação dos princípios de justiça é estabelecida através da concepção do justo e do bem, das liberdades políticas e de pensamento, por exemplo, Rawls fornece um status epistêmico para a crença nas convicções ponderadas em relação à eleição dos princípios de justiça.

No entanto, no processo da elegibilidade dos princípios de justiça, a concepção de pessoa, também, tem uma relevância teórica; isto, porém, não significa um *a priori* fundacionista, porquanto no construtivismo rawlsiano a idéia da concepção de pessoa é política; apenas um ideal razoável, onde este reflete ideais implícitos na cultura política pública das sociedades democráticas. Assim, a teoria rawlsiana limita o alcance do ideal da autonomia da pessoa. Neste sentido, as partes só elegem os princípios de justiça que já se encontravam presentes como forma de opção existente para eles; os princípios, por sua vez, só são considerados justos por causa das características pertinentes as partes que os escolheram.

Assim sendo, não só na elegibilidade dos princípios de justiça e, portanto na síntese efetuada por Rawls entre a liberdade dos antigos e dos modernos como, também, em todo o aparato teórico rawlsiano, por utilizar o método do equilíbrio reflexivo e, por conseguinte, ser equitativa, o que se constata é a presença do caráter holístico da justiça como equidade, corroborando com a tese de que a justificação do pensamento rawlsiano é coerentista.

REFERÊNCIAS

- BONELLA, Alcino E. (2000). *Justiça como imparcialidade e contratualismo*. Campinas, IFCH- Unicamp (Tese de Doutorado).
- BORON, Atilio A. *Filosofia política moderna: de Hobbes a Marx*. [on line] <http://bibliotecavirtual.clacso.org.ar/ar/libros/secret/filopolmpt/filopolmpt.html>
- BRINK, David O. (1989). *Moral Realism and The Foundations of Ethics*. New York: Cambridge University Press.
- CONSTANT, Benjamin. (1985). “Da liberdade dos antigos comparada à liberdade dos modernos”. *In: Filosofia Política*. Porto Alegre: LPM Editores.
- DANIELS, Norman. (1996). *Justice and Justification: Reflective Equilibrium in Theory and Practice*, New York: Cambridge University Press.
- GETTIER, Edmund. (1963). “Is Justified True Belief Knowledge?” *In: Analysis* No. 23.
- GONDIM, Elnora & MARRA, Osvaldino. (2008). “Rawls e a herança de Hobbes: convergências e divergências – um esboço”. *In: Revista Intuitio*, vol. I, no. 1, Porto Alegre.
- KYMLICKA, Will. (2006). *Filosofia Política Contemporânea*. São Paulo: Martins Fontes.
- LOCKE, John. (1994). “Segundo Tratado Sobre o Governo Civil: e outros escritos”. *In: Clássicos do Pensamento Político*. Petrópolis: Vozes
- OLIVEIRA, Manfredo. (1993). *Ética e Sociabilidade*. São Paulo: Loyola.
- OLIVEIRA, Nythamar Fernandes. (1999). “Hobbes, Liberalismo e Contratualismo”. *In: id. Tractatus ethico-politicus*. Porto Alegre: Edipucrs.
- _____. (2003). *Rawls*. Rio de Janeiro: Zahar.
- O’NEIL, Onora. (2006). *Em direção à Justiça e à Virtude: uma exposição construtiva do raciocínio prático*. São Leopoldo: Unisinos.
- RAWLS, John. (1997). *Uma Teoria da Justiça*. São Paulo: Martins Fontes.
- _____. (2003). *Justiça como Equidade: uma reformulação*. São Paulo: Martins Fontes.
- _____. (1998). *Justiça e Democracia*. São Paulo: Martins Fontes.
- _____. (2000). *O Liberalismo Político*. São Paulo: Ática.
- _____. (1999). *Collected Papers*.(org. Samuel Freeman). Cambridge: Harvard University Press.
- RICOEUR, Paul. (1997). *O Justo ou a Essência da Justiça*. Lisboa: Instituto Piaget.
- ROUSSEAU. (1978). “Do contrato social”. *In: Coleção Os Pensadores*. São Paulo: Abril Cultural.
- SILVEIRA, Denis Coitinho. (2007). “Teoria da justiça de John Rawls: entre o liberalismo e o comunitarismo”. *In: Trans/Form/Ação*, vol. 30, no. 1, p. 169-190.

CRÍTICAS AL ENFOQUE CAUSAL DE LA ACCIÓN PROPUESTO POR HARRY FRANKFURT*

CRITICISMS TO THE CAUSAL APPROACH OF THE ACTION PROPOSED BY HARRY FRANKFURT

JORGE GREGORIO POSADA RAMÍREZ
Universidad del Quindío, Colombia. gposada@uniquindio.edu.co

RECIBIDO EL 1 DE SEPTIEMBRE DE 2010 Y APROBADO EL 30 DE SEPTIEMBRE DE 2010

RESUMEN ABSTRACT

En este texto intento mostrar que la crítica de Frankfurt a lo que él denomina enfoque historicista causal de la acción es injustificada. Una teoría de la acción que distinga a las acciones de los sucesos a partir de su historia causal, no implica que la explicación de la acción se desplace a algo que ocurrió antes de la acción, su causa, o que descuide en la explicación de lo que es una acción, la acción misma. Para esto, muestro cómo la posición de Davidson y Searle, quienes son asumidos por Frankfurt como historicistas causales, si bien parten de la noción de causa como esencial para explicar lo que es una acción, no suponen que debe repararse en algo que sucede antes de la acción para explicarla, o que la causa de una acción sea un suceso que necesariamente la antecede.

My intention in this paper is to show that the Frankfurt's critique of what he determines causal historicism approach of the action is unjustified. A theory of the action that distinguishes the actions of the events from their causal history, does not mean that the explanation of the action moves itself to something that happened before the action, or that the action by itself, has to be neglected in its explanation. To demonstrate this, I show how the positions of Davidson and Searle, who are assumed by Frankfurt as causal historicist, although they are based on the notion of cause as essential to explain what an action is, it does not imply that we have to focus on something that happens before an action to explain it, or that the cause of an action is an event that necessarily precedes it.

PALABRAS CLAVE KEY WORDS

acciones, causas, Davidson, Frankfurt, Searle.

actions, causes, Davidson, Frankfurt, Searle.

* Las ideas que este artículo analiza hacen parte de algunos de los resultados de la investigación: *Ontología de la realidad social, una perspectiva de investigación para las ciencias sociales*, proyecto de investigación aprobado por la Universidad del Quindío.

El problema de la acción es inaugurado por Aristóteles cuando trata de distinguir entre aquellos acontecimientos que simplemente le suceden al hombre y aquellos que se dan gracias a que él es voluntariamente su principio generador. Aristóteles caracteriza a las acciones humanas como aquellas en donde los agentes son su principio causal: “como los hijos son de sus padres”¹. Austin en *Un alegato en pro de las excusas* afirma que una adecuada comprensión de la responsabilidad moral, debe estar antecedida por un análisis de lo que significa que el hombre hace algo: “Sin embargo, antes de considerar qué acciones son buenas o malas, justas o injustas, lo apropiado es considerar primero lo que se quiere decir con, y lo que no se quiere decir, y lo que se incluye en, y lo que no se incluye en, la expresión ‘realizar una acción’, o ‘hacer algo’”². Algunas de las teorías filosóficas que asumen este problema responden que en la causa de los sucesos se puede encontrar el rasgo distintivo. Un acontecimiento en el mundo es una acción si y sólo si es el resultado o puede asumirse como el resultado de cierto tipo de antecedentes. El filósofo de la mente y de la acción Harry Frankfurt define esta posición como: *el enfoque causal de la acción* y muestra que de manera similar a como lo plantea Aristóteles, el enfoque causal asume que para distinguir entre acciones y meros sucesos es necesario desplazar el análisis de la acción o del movimiento corporal a su génesis. Y si bien, la caracterización de lo que es la causa de una acción puede variar, marcando puntos divergentes dentro de los que asumen el enfoque causal, por ejemplo, Donald Davidson, John Searle, David Pears, el rasgo común de estos tipos de teoría es que asumen que las acciones son un tipo especial de efectos que son definidos y distinguidos de los meros sucesos por sus causas: “una acción consiste en producir, pero también evitar, intencionalmente, un suceso o cambio en el mundo. Más precisamente, una acción puede describirse apropiadamente como provocar o efectuar (a voluntad) un cambio”³.

Como casos palmarios, estornudar, sudar, roncar en tanto causados por disposiciones fisiológicas, son sucesos que les pasa a las personas. Servir una taza de café, disparar un arma, meter un gol son cosas que la gente hace, y que se puede explicar, asumiendo la intención o el querer de las personas de “servir una taza de café”, “disparar una bala” o “meter un gol”, como causas de estos acontecimientos.

¹ ARISTÓTELES. *Ética Nicomaquea*. Libro 3, Capítulo 5.

² AUSTIN, John. *Ensayos filosóficos*. Madrid: Biblioteca de la Revista de Occidente, 1975. p. 171.

³ VON WRIGHT, George. *Sobre la libertad humana*. Barcelona: Paidós, 2002. p. 17.

Para Frankfurt, la motivación más importante del enfoque causal lo constituye el presupuesto de que los sucesos y las acciones son intrínsecamente indistinguibles. Si se considera a las acciones, sin su historia causal, únicamente como efectos, obviando sus causas, no puede decidirse cuándo se está ante una acción o cuándo ante mero suceso. Esto porque no hay una característica inherente en la acción que permita diferenciarla de los movimientos corporales. Un mismo movimiento corporal puede ser o no ser una acción. El movimiento de la cabeza de alguien cuando está asintiendo, vale algunas veces como acciones y otra como un mero suceso. Los movimientos reflejos de una persona que guiña su ojo por un tic nervioso, en nada se distinguen de los de un hombre que lo hace siguiendo sus deseos de agradar. La distinción aparece sólo cuando se repara en las secuencias causales que originan los movimientos. Si la persona que guiña su ojo lo hace motivada por el deseo y las creencias de que así conseguirá dar señales de agrado, es un hombre que ejecuta una acción. Así, la diferencia entre la acción de asentir y de guiñar el ojo, con la de los mismos movimientos corporales, pero en los que no se asiente ni se busca agradar, está en los orígenes de ambos movimientos. Los movimientos del asentir y el guiñar son causados por deseos y creencias, por esto son acciones, mientras que los de los simples sucesos, en tanto causados por actos reflejos, por ejemplo, no son acciones. Al respecto Frankfurt señala:

Cuando afirman que la diferencia esencial entre las acciones y los meros sucesos reside en sus historias causales previas, las teorías causales presuponen que en esencia las acciones y los meros sucesos no difieren. Estas teorías sostienen que las secuencias causales que producen acciones necesariamente son de un tipo diferente de aquellas que producen meros sucesos, pero que los efectos producidos por las secuencias de ambos tipos son, desde un punto de vista intrínseco, indistinguibles⁴.

El supuesto del enfoque causal de que no hay nada inherente en las acciones que las permita distinguir de los meros sucesos o movimientos corporales, induce a los filósofos de esta perspectiva a desplazar, el análisis de lo que sea una acción, de la acción misma al suceso que la antecede. Así, para saber cuando algo es una acción, el análisis debe trasladarse de la acción misma al suceso que la provoca: su causa. No obstante, como afirma Frankfurt, parece contra intuitivo afirmar que si

⁴ FRANKFURT, Harry. La importancia de lo que nos preocupa. Buenos Aires: Katz, 2006. p. 106.

una persona sabe que está haciendo algo su conocimiento no se debe al reconocimiento de lo que está haciendo en ese momento, sino que lo debe de haber obtenido de ciertas condiciones previas, que ya pasaron.

Esto hace que las teorías causales sean poco razonables. Alejan la atención exclusivamente de los acontecimientos cuya naturaleza se analiza y de los momentos en que ocurren. El resultado es que queda más allá de su alcance estipular que la persona debe relacionarse de algún modo específico con los movimientos de su cuerpo durante el período en que se presume que está llevando a cabo una acción⁵.

En su ensayo *El problema de acción* Frankfurt hace una crítica a los que se circunscriben en el enfoque causal, entre ellos Donald Davidson y John Searle, en virtud a que estas teorías (i) desplazan su análisis de lo realmente relevante cuando ocurre la acción, la acción misma, y no sus antecedentes, y (ii) porque no pueden explicar los contraejemplos de aquellos sucesos que ocurren, a pesar de que tengan como origen las condiciones causales necesarias y suficientes que suponen estas teorías son prototípicos de la acción. Para Frankfurt, las teorías causales no pueden explicar casos del tipo Chisholm, casos en los que se da la historia causal necesaria para provocar la acción, se da el estado de cosas en el mundo que se supone es causado por la historia causal relevante, pero realmente no se tiene una acción⁶. Por estas razones para Frankfurt la posición de Davidson y la de Searle, en tanto basadas en el enfoque causal dejan demasiado por explicar a la pregunta qué es realizar una acción o hacer algo, de ahí que deben ser abandonadas por otro enfoque más promisorio⁷.

⁵ *Ibidem*.

⁶ Un arquero de fútbol ha estudiado meticulosamente la manera como patea los penaltis un jugador contrario. Sabe que para llevar a cabo la acción de tajar la pena máxima, no debe lanzarse para ningún lado, sino quedarse en el centro del arco, no obstante, a la hora de enfrentarse con la situación está tan asustado, por lo que significa para su país tajar el penalti, que sus nervios lo dejan helado en el centro de su arco. El arquero despierta de su pasmo cuando el balón pega en su cuerpo evitando que entre en el arco. Piensa Frankfurt que casos de este tipo no pueden ser explicados desde el enfoque causal, pues se supone que se dan las condiciones causales relevantes, se dan en el mundo los estados de cosas que deseaba el agente, pero no parece que esto sea algo que realmente hace el agente.

⁷ Frankfurt sugiere que un rasgo común a las acciones podría ser el que son movimientos guiados. Las acciones humanas registran patrones de movimiento, detectarlos podría servir para distinguirlas de los meros sucesos, y dado que ocurren en el momento en el que la persona está realizando la acción, esta alternativa podría evitar (i). Los movimientos de las manos de un pianista registran patrones de movimientos, son acciones, los de un hombre que está sufriendo un ataque epiléptico son meros sucesos. No obstante, el mismo Frankfurt reconoce que otros problemas podrían dejar mal librada esta alternativa. Cómo distinguir patrones de movimientos en acciones básicas. La sugerencia de Frankfurt podría arrastrar al enfoque causal, pues un patrón de movimiento bien podría ser entendido como movimientos causalmente guiados por agentes.

En lo que sigue intentaré mostrar que ni Davidson ni Searle, a pesar de que asumen como relevante la noción de causa para explicar lo que son las acciones, siguen los supuestos que Frankfurt considera necesarios en el enfoque causal. Mostraré que el hecho de describir una acción a partir de su causa, no implica que tenga que apelarse a un suceso distinto e independiente de la acción; pues las causas no siempre son cosas distintas, independientes y que ocurren antes que sus efectos. De ahí que las críticas de Frankfurt a lo que él denomina el enfoque causal, bien pueden ser una falacia del hombre de paja.

RAZONES COMO EXPLICACIONES CAUSALES DE LAS ACCIONES

El clásico ensayo de Donald Davidson, *Acciones razones y causas* de 1963, es para Frankfurt, dentro del modelo causal, la descripción prototípica. En él Davidson afirma que un rasgo distintivo de una acción, en contraste con un mero suceso, consiste en que en la acción puede haber una razón que explica causalmente la acción. Siguiendo la posición clásica, Davidson define racionalización a la explicación causal que señala dentro de una razón algo que el agente quiso, apreció, deseó, que le pareció atractivo, benéfico, obligatorio. En tanto se pueda explicar un acontecimiento mostrando un rasgo de este tipo, para Davidson el acontecimiento puede entenderse como una acción racionalizada, pues se describe apelando a su razón primaria. En *Acciones, razones y causas* Davidson dice:

A menudo, dar la razón por la cual un agente hizo algo consiste en nombrar la actitud favorable (a) o la creencia relacionada (b), o ambas; permítaseme llamar a este par la *razón primaria* por la que el agente realizó la acción. Ahora es posible reformular la afirmación de que las racionalizaciones son explicaciones causales y también dotar de una estructura al argumento enunciando dos tesis sobre las razones primarias:

1. Para entender cómo una razón de cualquier tipo racionaliza una acción, es necesario y suficiente que veamos, por lo menos en sus rasgos esenciales, cómo construir una razón primaria.
2. La razón primaria de una acción es su causa⁸.

⁸ DAVIDSON, Donald. Ensayos sobre acciones y sucesos. Barcelona: Crítica, 1995. p.18.

Provisionalmente puede decirse que un acontecimiento es una acción y no un mero suceso, si puede mostrarse la causa de ese acontecimiento como una razón primaria. En tanto se describa la génesis de un acontecimiento como una razón primaria no tenemos un mero suceso sino una acción, en caso contrario, se tiene un mero suceso. No obstante, la interpretación de Frankfurt da un paso más allá. Frankfurt sugiere que Davidson, como todos los historicistas causales, en tanto postula la razón primaria de un acción como su causa, asume que la razón primaria provoca diferencias intrínsecas entre las acciones y los meros sucesos. Pero en Davidson la afirmación: “la razón primaria de una acción es su causa”, no implica que asuma características ontológicas distintivas entre acciones y sucesos. Esto se hace claro en el ejemplo con el que Davidson representa la relación de la razón primaria con la acción. “Presiono el interruptor, enciendo la luz e ilumino el cuarto. Sin saberlo, también alerto a un merodeador de mi presencia en la casa”⁹. Muestra Davidson que son cuatro descripciones distintas de un mismo suceso. Y si bien, puede asumirse que la acción ejecutada es, por ejemplo, “presionar el interruptor”, – dado que se tiene la razón primaria “quería encender la luz” – a pesar de que no haya una razón primaria que explique “alerté a un merodeador” no obstante, es algo que también se hizo. Es una acción aún no racionalizada o explicada por una razón primaria.

Las razones primarias racionalizan las acciones pero no las caracterizan intrínsecamente de los meros sucesos. Puede darse el caso de que se den acciones sin que necesariamente se tenga la razón primaria que las racionaliza. Esto se confirma si partimos de la afirmación de Davidson de que dar una razón primaria es dar a conocer la intención por la cual un agente hizo lo que hizo. Conocer la razón primaria de una acción es conocer la intención de la acción. No obstante, toda acción no necesariamente es intencional, ni exige una razón primaria que la racionalice. No hay una razón primaria en la acción de alertar al merodeador, ni una intención, pues no puede verse en el agente un deseo o una actitud favorable de alertar al merodeador, sin embargo es una acción, es una cosa que se dice hace el agente. Así, hay cosas que se hacen no intencionalmente y son acciones, que no necesariamente son racionalizadas pero son cosas que hace la gente. El defensa de un equipo de fútbol a pesar de que no tenga la intención de meter en su propio arco un gol, si lo hace, hace una acción, la de meter un autogol. Davidson dice: “Así, derramar el café, hundir el Bismarck, e insultar

⁹ *Ibid.*, p. 19.

a alguien son todas cosas que pueden o no hacerse intencionalmente, pero incluso cuando no son intencionales, son normalmente acciones”¹⁰.

La posición de Davidson es que ningún acontecimiento es realmente intencional. Lo que hay son descripciones intencionales. La fórmula de Davidson es que si puede encontrarse una descripción de un suceso que haga verdadera una oración que dice que se hizo intencionalmente, tenemos la descripción de una acción. No obstante, y retomando lo que nos interesa, puede afirmarse que no es a partir de lo que entiende Davidson como razones primarias que puede vincularse como defensor de lo que Frankfurt entiende como el enfoque causal. Recordemos que para Frankfurt este enfoque distingue las acciones de los meros sucesos a partir de sus antecedentes causales, los cuales son sucesos que se dan antes de que se den las acciones; en tanto Davidson no asume las razones primarias como sucesos ontológicamente reales que anteceden causalmente a las acciones, sino como la racionalización de una acción, y esto es, como una descripción que muestre una actitud favorable de un agente, y no un suceso distinto, e independiente que antecede a otro, entonces, no es desde este punto de vista que puede asumirse como un defensor de lo que Frankfurt considera el enfoque causal, ni endosársele los problemas que para Frankfurt este enfoque presenta. Las razones primarias son descripciones causales que racionalizan una acción, no sucesos o estados de cosas que causan o producen una acción.

Cinco años después de publicar *Acciones, razones y causas*, Davidson presenta su ensayo *De la acción*. La pregunta ya no es cómo una razón racionaliza o describe causalmente la explicación de una acción, sino cómo distinguir las acciones que ejecuta un agente de los meros sucesos que le pasan. Davidson cree que en este caso sí es clave la noción de causa. Pero de nuevo vuelve a enfatizar que es en el nivel de la descripción que puede afirmarse que una acción ha sido ejecutada por un agente, no en el sentido de una cosa que es, por ejemplo, intencional causa otra cosa que sería una acción.

Como un primer paso para enderezar las cosas podemos intentar hablar, no de acciones, sino, en su lugar, de oraciones y descripciones de acciones. En el caso de la actuación, mi propuesta podría entonces formularse así: una persona es el agente de un suceso si y sólo si hay una descripción de lo que ella hizo que haga verdadera una oración que dice que lo hizo intencionalmente¹¹.

¹⁰ *Ibid.*, p. 65.

¹¹ *Ibid.*, p. 66-67.

Así, la afirmación de Frankfurt según la cual, la posición causal de Davidson está equivocada porque su enfoque causal implica “que no puede haber una acción sin intencionalidad o cuando no se cumplen las condiciones actitudinales pertinentes”¹² obvia la afirmación de Davidson de que puede haber acciones sin intenciones, cosas que un agente hace, sin que se sepa la razón que racionaliza su acción; pues si bien las “intenciones implican acciones”, la inversa no es verdadera: las acciones no necesariamente implican intenciones, por lo menos como sucesos del mundo que anteceden causalmente a las acciones.

LA INTENCIÓN COMO CAUSA DE LA ACCIÓN

En su análisis de las acciones Davidson no se compromete con ninguna afirmación que lleve a pensar que una acción es el efecto de un suceso o estado distinto de la acción. El hecho de que se describan ciertos sucesos ‘como si’ fueran acciones, y otros ‘como si’ fueran causas intencionales de esas acciones, no implica que para Davidson de hecho haya acciones y causas intencionales de esas acciones. La teoría de Davidson al concentrarse en el nivel de la descripción, y no en el de los sucesos mismos, se mantiene al margen de compromisos causales reales entre sucesos y acciones. Esto basta para mostrar que Frankfurt se equivoca en su crítica a Davidson, al considerarlo defensor de lo que define como enfoque causal. No obstante, puede pensarse que una teoría que sea realista con respecto a las intenciones como estados que se dan en el mundo, y a las acciones, como consecuencias de estos estados, sí pueda verse mal librada de los ataques de Frankfurt. En lo que sigue intentaré mostrar que a pesar de que la teoría de la acción de John Searle es realista, en el sentido de afirmar que hay estados intencionales y acciones como efectos de esos estados, la crítica de Frankfurt que las teorías causales desplazan su explicación de lo que sea una acción a un suceso anterior a la acción, su causa, y que por lo tanto, descuidan lo que es relevante en la explicación de la acción: la acción misma, no es correcta; por lo menos en el planteamiento de Searle. El reproche para Frankfurt es que siendo la teoría causal de la acción de Searle un caso representativo de enfoque causal de la acción, si los problemas que supone Frankfurt arrastran las teorías historicistas causales de la acción, no son tales en la propuesta de Searle, las críticas de Frankfurt van dirigidas a un enfoque que él mismo ha construido: el enfoque historicista causal, para luego derribarlo.

¹² FRANKFURT, La importancia de lo que nos preocupa. Op. cit., p. 115.

En su ensayo sobre filosofía de la mente, *Intencionalidad*, John Searle describe y argumenta a favor de la causación mental. Existen estados mentales que causan ciertas cosas en el mundo, y existen estados de cosas en el mundo que tienen poder causal en los estados mentales de las personas. Las acciones son para Searle un hecho palmario de esto: estados mentales que causan ciertas cosas en el mundo. Por eso cuando describe lo que es una acción Searle afirma que dos cosas tienen que darse necesariamente para que haya una acción: movimientos físicos y estados mentales intencionales que causen dichos movimientos físicos. Esto es consistente con la crítica de Frankfurt, pues se da una acción cuando un estado causa que se dé otro estado en el mundo. La misma noción de causalidad que Searle propone: “la causación es una relación real entre objetos y acontecimientos del mundo, una relación por la que un fenómeno, la causa, causa otro, el efecto”¹³, encaja consistentemente con la idea que Frankfurt considera defectuosa en la explicación de las acciones de los historicistas causales. Como ya se mencionó, en tanto puede darse cierto tipo de movimientos físicos, que son meros sucesos, pero que son idénticos a los movimientos que son acciones –por ejemplo, el movimiento físico de asentir o de guiñar el ojo–, el movimiento físico es necesario pero no suficiente para que se dé una acción, y precario para distinguirlo de los meros sucesos. En Searle la otra condición necesaria para la acción es el estado intencional relevante como causa del movimiento físico.

Lo anterior puede empujar a la creencia de que Searle desplaza la explicación de lo que sea una acción, de las cosas que suceden en el momento en el que se da la acción, a lo que es su causa: el estado mental intencional, el cual, se supone desde el enfoque causal de Frankfurt, ha de haber ocurrido antes de que se dé el movimiento corporal, dado que es su causa. No obstante, esto es ir demasiado ligero con el enfoque de Searle sobre la acción y desconocer en buena medida su teoría. Pues, si bien John Searle cree que algunas acciones tienen como causas estados mentales previos a los movimientos físicos, también afirma que esto se da sólo algunas veces. Searle distingue entre intenciones que son formadas con anterioridad a las acciones e intenciones que son formadas en el mismo momento en que son llevadas las acciones: intenciones previas e intención en la acción, respectivamente. Si bien la acción de salir de la ciudad estuvo antecedida por la intención previa de salir de la ciudad, esta acción, para Searle, está subsidiada por otras acciones

¹³ SEARLE, John. *Mente, lenguaje y sociedad*. Madrid: Alianza, 2004. p. 21.

que son llevadas a cabo como parte de la acción de salir de la ciudad, pero que en sí mismas, no son las acciones de salir de la ciudad. Abrir la puerta, bajar las escaleras, entrar al parqueadero, encender el carro, entre muchas otras, son acciones que subsidian la acción de salir de la ciudad, no obstante, no son necesarios estados intencionales previos a estas acciones para que estas acciones se den. Estos son casos de intención en la acción.

Hay al menos dos modos de establecer más claramente la distinción entre una intención en la acción y una intención previa. El primero es observar que muchas de las acciones que se realizan, se realizan bastante espontáneamente, sin formar consciente o inconscientemente, ninguna intención previa de hacer esas cosas. Por ejemplo, supongamos que estoy sentado en una silla reflexionando sobre un problema filosófico, y de repente me levanto y comienzo a pasearme con aspecto preocupado por la habitación. Mi levantarme y pasear con aspecto preocupado son claramente acciones intencionales, pero para hacerlas no necesito formar una intención para hacerlas antes de hacerlas (...) Como muchas de las cosas que se hacen, yo solamente hago estas acciones; solamente actúo¹⁴.

Para Searle, esa serie de acciones que corrientemente se realizan espontáneamente, en tanto acciones, es causada por estados mentales intencionales, pero estados mentales intencionales que no son previos, sino simultáneos a la acción; él los denomina intención en la acción, intención que si bien determina la experiencia de actuar, no es un acontecimiento discreto y distinto de la acción misma. De manera similar, Davidson afirma que hay acciones en las que el agente simplemente actúa, sin que esto implique una causa antecedente previa a la actuación. Davidson las llama acciones primitivas "hacer algo que cause que mi dedo se mueva no me causa que yo mueva mi dedo, es mover mi dedo"¹⁵.

La descripción de Searle de lo que es la experiencia de actuar, apoya la tesis de que no es necesario un antecedente previo para que se lleve a cabo una acción. Con dos hechos psicológicos muestra que en la descripción de lo que es una acción, es condición necesaria además del

¹⁴ SEARLE, John. Intencionalidad, un ensayo en la filosofía de la mente. Madrid: Tecnos, 1992. p. 96.

¹⁵ DAVIDSON, Op. cit., 71.

movimiento corporal, una experiencia de actuar, y ésta es simultánea a la acción causándola. Así, su teoría no implica, lo que Frankfurt supone implica todo enfoque causal de la acción: “Para el enfoque causal, es esencial considerar que las acciones y los meros sucesos no se diferencian por nada que existe o que sucede en el momento en que estos acontecimientos ocurren, sino por algo que les es extrínseco: una diferencia, en un momento previo, en otra serie de acontecimientos”¹⁶. Las experiencias fenoménicas que tienen consecuencias lógicas para la explicación de la acción, pues implican como condición necesaria de la acción la experiencia de actuar, las describe Searle a partir de los casos que toma de William James y de Penfield, de ellos se concluye que pueda darse experiencia de actuar pero no movimientos físicos, o movimientos físicos pero no experiencia de actuar, por lo que la experiencia de actuar y el movimiento físico son condiciones lógicamente necesarias para la acción.

En el caso de James, citado por Searle, un paciente con un brazo anestesiado es inducido a que mueva su brazo. El paciente no puede ver que su brazo está atado, así que a la petición de que levante su brazo asiente, pero a pesar de que tiene la experiencia de levantar el brazo, realmente no ejecutó la acción de levantar el brazo. Searle concluye que no basta con el estado intencional de levantar el brazo, sino que de hecho tiene que darse el movimiento físico para que haya una acción. No obstante, el caso muestra que es un hecho que la experiencia de actuar es parte constitutiva de la acción. Cuando se tiene una acción, además del movimiento físico, es necesaria la experiencia de que ese movimiento físico es causado por un agente. Los casos Penfield se suman a la justificación de lo anterior. Un doctor estimulaba a través de electrodos el córtex motor de uno de los hemisferios cerebrales de un paciente, causando el movimiento de una de sus manos. El paciente siempre afirmaba que él no había movido su mano, que eso lo había hecho el doctor. En este caso, distinto al de James, se da el movimiento físico, pero no se da la experiencia de que ese movimiento físico ha sido causado por el agente. Faltó lo que Searle denomina la experiencia de actuar. Así, la experiencia psicológica que percibe el agente que ejecuta la acción, sumada al movimiento físico pertinente con la experiencia de actuar constituye la acción. La experiencia de actuar y el movimiento físico son para Searle las condiciones necesarias para la constitución de una acción.

¹⁶ FRANKFURT, La importancia de lo que nos preocupa. Op. cit., p. 106.

Para Searle, a diferencia de lo que supone Frankfurt de los enfoques causales, las acciones son casos de causación simultánea. En la filosofía de la mente de Searle es recurrente la alusión de relaciones causales que no implican un antes y un después de dos eventos o sistemas distintos: la causa y el efecto. Searle anota "(...) exactamente en este sentido de "causado por" y "realizado en" los fenómenos mentales son causados por procesos que tienen lugar en el cerebro en el nivel neuronal o modular, y al mismo tiempo se realizan en el sistema mismo que consta de neuronas"¹⁷. Argumenta que aceptar que las relaciones causales siempre implican un antes y un después es matricularse con una tesis rudimentaria de la causalidad, ya superada por hechos científicos. Las partículas de H₂O por sí solas no son líquidas, pero cuando hay un buen número de ellas causan simultáneamente la liquidez. Las acciones, entendidas como la conexión simultánea y causal entre la experiencia de actuar, su causa, y el movimiento físico, su efecto, son efectivamente sucesos que se distinguen de otros sucesos por sus causas; pero de esto no se sigue, como lo sugiere Frankfurt, que exista algo detrás, intrínsecamente distinto de las acciones mismas, su causa. Pues una acción es la experiencia de actuar y el movimiento físico, y a pesar de que la primera cause la segunda, no son eventos completamente distintos, en donde uno está antes y el otro está después.

Así, si tanto la teoría de Davidson como la de Searle son prototípicas del enfoque historicista causal de la acción que describe Frankfurt, pero ninguna presupone lo que Frankfurt afirma conviene a toda teoría causal de la acción, la crítica de Frankfurt se dirige a un enfoque que él mismo construye para desarmarlo.

REFERENCIAS

- ARISTÓTELES. (1967). Obras completas. Madrid: Aguilar.
- AUSTIN, John. (1975). Ensayos filosóficos. Madrid: Biblioteca de la Revista de Occidente.
- DAVIDSON, Donald. (1995). Ensayos sobre acciones y sucesos. Barcelona: Crítica.
- FRANKFURT, Harry. (1975). "Three concepts of free actions". En: Aristotelian Society, vol. 49, p. 95-125.
- _____. (2006). La importancia de lo que nos preocupa. Buenos Aires: Katz.

¹⁷ SEARLE, John. Mentes cerebros y ciencia. Madrid: Cátedra, 2001. p. 27.

- _____. (2007). Necesidad, volición y amor. Buenos Aires: Katz.
- HUME, David (1995). Investigación sobre el conocimiento humano. Barcelona: Altaya.
- POSADA, Gregorio (2007). La noción tripartita del conocimiento, una introducción a la epistemología. Manizales: Universidad de Caldas.
- _____. (2007). "Una aproximación al problema de la acción". En: Jurídicas, vol. 4, p. 79 -91.
- SEARLE, John. (1979). "What is an intentional State?" En: Mind, New Series. Oxford, vol. 88, p. 74-92.
- _____. (1992). Intencionalidad, un ensayo en la filosofía de la mente. Madrid: Tecnos.
- _____. (2001). Mentes cerebros y ciencia. Madrid: Cátedra.
- _____. (2004). Mente, lenguaje y sociedad. Madrid: Alianza.
- RUSSELL, Bertrand. (1950). Análisis del espíritu. Buenos Aires: Paidós.
- VON WRIGHT, George (1979). Explicación y comprensión. Madrid: Alianza.
- _____. (2002). Sobre la libertad humana. Barcelona: Paidós.

PRAGMÁTICA SIN PRAGMATISMO

PRAGMATICS WITHOUT PRAGMATISM

GABRIEL MÉNDEZ HINCAPIÉ

Universidad de Caldas, Colombia. gabriel.mendez@ucaldas.edu.co

RECIBIDO EL 31 DE AGOSTO DE 2009 Y APROBADO EL 30 DE SEPTIEMBRE DE 2010

RESUMEN ABSTRACT

Este artículo muestra la manera en que se relacionan algunos aspectos pragmáticos en la obra de Wittgenstein con elementos centrales del pragmaticismo de Ch. S. Peirce. En él se opta por una vía alternativa que supera el dualismo wittgensteiniano al partir de un esquema semiótico-modal genuinamente triádico, es decir, no-dualista. En particular, se pretende mostrar que dos herramientas auténticamente triádicas, a saber, los *gráficos existenciales* y la *semiótica universal* de Charles S. Peirce, integradas en un modelo de "lógica topológica", constituyen un instrumental válido y original para re-dimensionalizar la noción de 'juego de lenguaje'.

This article shows how relate some pragmatic aspects in the work of Wittgenstein to core elements of Ch. S. Peirce pragmaticism . It chooses an alternative that overcomes the dualism wittgensteinian from an outline genuinely triadic semiotic-modal, ie, non-dualistic. In particular, it aims to show that genuinely triadic two tools, namely the existential graphs and universal semiotics of Charles S. Peirce, integrated into a model of "topological logic" constitute a valid and original tools in order to re-dimension the notion of 'language-game'.

PALABRAS CLAVE KEY WORDS

dualismo, gráficos existenciales, juego de lenguaje, lógica topológica, pragmaticismo, Semiótica triádica.

dualism, existential graphs, language-game, topological logic, pragmaticism, Triadic semiotics.

En *Los cuadernos azul y marrón*, Ludwig Wittgenstein introduce la noción de ‘juego de lenguaje’ con la intención de elaborar una teoría capaz de explicar lo que sucede en el mundo al ser representado por “sistemas completos de comunicación humana”¹. Sin embargo, al estar incluida en un esquema lógico que implícitamente supone el dualismo², la noción suscita algunas consecuencias restrictivas; en especial, queda descartada como sinsentido toda argumentación metafísica.

Este artículo opta por una vía alternativa y propone partir de un esquema semiótico-modal genuinamente triádico, es decir, no-dualista. En particular, se pretende mostrar que dos herramientas auténticamente triádicas, a saber, los *gráficos existenciales* y la *semiótica universal* de Charles S. Peirce, integradas en un modelo de “lógica topológica”, constituyen un instrumental válido y original para re-dimensionar la noción de ‘juego de lenguaje’.

La importancia de esta re-dimensión reside en el hecho de que dicha noción es una re-formulación del asunto, según Kant, más complejo de la filosofía: “Kant señaló que el problema de cómo lo que hay en la mente puede ser una ‘representación’ de lo que hay fuera de ella es el asunto más complejo de la filosofía”³. Y como Wittgenstein es el filósofo más influyente del siglo XX, una lectura alternativa de su propuesta nos llevará hacia una nueva y más fructífera solución de dicho asunto.

A diferencia de lo que afirma Richard Rorty, no fueron Frege y Wittgenstein los que primero sustituyeron –en el denominado giro lingüístico de la filosofía– el problema de cómo lo que hay en la mente puede ser una “representación” por el de cómo se relaciona el lenguaje con el mundo. En realidad fue Peirce⁴ el que ya antes había reformulado la problemática kantiana en los términos de su propio sistema filosófico, el *pragmatismo*: “La verdad es la conformidad de una representación

¹ Cfr. WITTGENSTEIN, Ludwig. *Los cuadernos azul y marrón*. Madrid: Editorial Tecnos 1984, p. 115-116.

² “La diferencia esencial entre estos últimos Wittgenstein y Frege, y Peirce es que para Peirce la lógica no es solamente una teoría de la inferencia a partir de signos, sino de la inferencia a partir de signos mediante signos. Las teorías de Frege y Wittgenstein son diádicas y dualistas; la de Peirce es triádica y dialéctica”. DELEDALLE, G. *Leer a Peirce hoy*. Barcelona: Editorial Gedisa 1996. p. 163 (subrayado añadido).

³ PUTNAM, Hillary. *Cómo renovar la filosofía*. Madrid: Ediciones Cátedra, 1994. p. 55

⁴ De hecho el mismo Rorty le hace cierta justicia a Peirce cuando afirma que la expresión *man's giassy essence* fue mencionada por primera vez en filosofía por C. S. Peirce, en un trabajo en donde se plantea la opinión de que “una persona no es nada más que un símbolo que implica una idea general”.

con su objeto' dice Kant; se podría hacer más explícita esta afirmación (...) es casi correcta, en la medida en que es inteligible. Solo que, ¿qué es ese 'objeto' que sirve para definir la verdad? Es lo real"⁵.

Y qué es lo real para Peirce:

'Real' es una palabra inventada en el siglo trece, que significa tener Propiedades, esto es, características que basten para identificar a su objeto, y poseerlas ya le sean atribuidas o no por algún hombre singular o grupo de hombres (...) Constituyen [las propiedades] una serie de circunstancias suficientes para distinguirlo de todos los demás sucesos, y estas circunstancias le pertenecen, esto es, sería verdadero predicarlas de él, si A, B, o C las descubren actualmente o no. Lo 'Actual' es aquello que se tiene en el pasado, presente o futuro⁶.

En ese sentido, como se ha subrayado, la reformulación pragmaticista del "asunto más complejo de la filosofía" que hace Peirce, tiene una clara tintura modal⁷. Pero, a su vez, ¿qué entiende Peirce por pragmaticismo?

Ya que he empleado la palabra *pragmaticismo*, y tendré ocasión de usarla alguna vez más, quizás puede ir bien explicarla. Hace alrededor de cuarenta años, mis estudios de Berkeley, Kant y otros me condujeron, después de convencerme a mi mismo de que todo pensamiento se lleva a cabo mediante Signos y de que la meditación toma la forma de un diálogo, de forma que resulta acertado hablar del "significado" de un concepto, a concluir que para adquirir pleno dominio de ese significado se requiere, en primer lugar, aprender a reconocer el concepto bajo cualquier disfraz, mediante una gran familiaridad con sus casos. Pero esto, después de todo, no implica ninguna

⁵ PEIRCE, C. S. *Collected Papers of Charles Sanders Peirce*. HARTSHORNE, Ch., WEISS, P. & BURKS, A. (Eds). Cambridge: Harvard University Press, 1931-1958. En adelante se nombrará CP. Las referencias a los CP se realizan de la manera usual así: CP 1.578, en este ejemplo envía al párrafo 578 del primer volumen de los *Collected Papers*.

⁶ PEIRCE, C. S. Un argumento olvidado en favor de la realidad de Dios. [En línea]: www.unav.es/gep./Peirce-esp.html (subrayado añadido).

⁷ "Allá por 1860, cuando aún no conocía nada de ningún filósofo alemán distinto de Kant, que había sido durante tres o cuatro años mi reverenciado maestro, me llamó mucho la atención cierta indicación de que la lista kantiana de categorías podría no ser sino una parte de un sistema más amplio de conceptos. Así, por ejemplo, las categorías de relación (...) no son sino otros tantos modos distintos de la necesidad, que es (junto a la posibilidad y la actualidad) una categoría de la modalidad". PEIRCE, C. S. *Escritos lógicos*. Madrid: Alianza Editorial, 1968. p. 81.

comprensión verdadera de él; por eso es un requisito posterior el que hagamos un análisis lógico abstracto de él en sus últimos elementos, o un análisis tan completo como podamos lograr. Pero incluso entonces podemos todavía estar sin una comprensión viva de él; y el único modo de completar nuestro conocimiento de su naturaleza es descubrir y reconocer justamente qué hábitos generales de conducta podría desarrollar razonablemente la creencia en la verdad del concepto (de cualquier materia y bajo cualesquiera circunstancias concebibles); es decir, qué hábitos resultarían al final de una consideración suficiente de tal verdad. Es necesario entender la palabra “conducta”, aquí, en su sentido más amplio. Si, por ejemplo, la predicación de un concepto dado nos llevara a admitir que una forma dada de razonamiento relativo a la materia de la que se afirmó que era válido, cuando en caso contrario no sería válido, el reconocimiento de ese efecto en nuestro razonamiento sería decididamente un hábito de conducta.

En 1871, en un *Club Metafísico* en Cambridge, Massachusetts, solía predicar este principio como un tipo de evangelio lógico, representando al método no formulado que siguió Berkeley, y en una conversación sobre él lo denominé “Pragmatismo” (...) Por supuesto, aquella doctrina no atrajo particular atención, ya que, como había advertido en mi frase de apertura, muy poca gente se preocupa por la lógica. Pero en 1897 el Profesor James replanteó el asunto, y lo transformó en una doctrina de filosofía, de la que me parecieron muy bien algunas de sus partes, mientras que otras partes más prominentes las consideré – y todavía las considero – como opuestas a la buena lógica (...) [Entonces], llegué yo a la conclusión de que mi pobre pequeña máxima debería ser llamada por otro nombre; y de acuerdo con eso, en abril de 1905 le di el nuevo nombre de *pragmaticism*⁸.

Específicamente, por un esquema semiótico-modal genuinamente triádico, entendemos aquí el estudio de lo que Peirce llamó el ‘tercer universo’ o poder activo que establece conexiones entre ‘objetos’ pertenecientes a los otros dos universos, es decir, el estudio de la terceridad (*thirdness*):

⁸ *Ibidem*.

El tercer universo comprende todo aquello cuyo ser consiste en un poder activo para establecer conexiones entre objetos diferentes, especialmente entre los objetos de los diferentes Universos. Tal es todo lo que es esencialmente un Signo – no el mero cuerpo de un Signo, que no es esencialmente tal, sino, por decir así, el Alma del Signo, que tiene su Ser en su poder de servir de intermediario entre su Objeto y una Mente. Tal es también una conciencia viva y tal es la vida, el poder de crecimiento de una planta. Tal es una constitución viva -un periódico diario, una gran fortuna, un “movimiento” social⁹.

Las siguientes páginas tratan un problema tan viejo como la filosofía: la relación entre lo real y los diversos dispositivos de representación. Debe señalarse que no se intenta cubrir todos los dispositivos involucrados ni contrastar el punto de vista que aquí se adopta con la posición de otros autores que también sostiene que el lenguaje – el dispositivo de representación privilegiado – juega un papel fundamental.

LAS RELACIONES ENTRE PEIRCE Y WITTGENSTEIN

En su completo estudio titulado *Scholarship on the relations between Ludwig Wittgenstein and Charles S. Peirce*, que a continuación se reseña ampliamente, Jaime Nubiola señala cómo hace ya cuarenta años Rorty “destacó la proximidad entre el giro lingüístico wittgensteniano” y la concepción filosófica de Peirce.

Según Nubiola, Rorty sostenía que “Peirce había afrontado y rechazado con cinco décadas de antelación el empirismo positivista y había desarrollado un conjunto de y un estilo de hacer filosofía muy parecido al de los filósofos contemporáneos continuadores del segundo Wittgenstein”¹⁰. De modo semejante, señala que en tiempos más cercanos, “Christopher Hookway ha insistido en que la mejor manera de comprender a Peirce es entenderlo como un filósofo analítico *avant la lettre*”¹¹.

⁹ *Ibidem*.

¹⁰ NUBIOLA, J. “Scholarship on the relations between Ludwig Wittgenstein and Charles S. Peirce”. En: ANGELELLI, I. & CERESO, M. (Eds.). *Proceedings of the III Symposium on History of Logic*. Berlin: Walter de Gruyter GmbH & Co., 1996. p. 281.

¹¹ *Ibidem*.

En efecto, en Peirce están presentes muchas de las características distintivas de la filosofía analítica tales como su énfasis en el análisis lógico y en el respeto por la coherencia e, incluso, una posible contrastación de las hipótesis filosóficas por medios científicos aceptados.

Por otra parte, “Karl-Otto Apel identifica el pensamiento de Peirce como la piedra de toque del proceso de transformación semiótica de la filosofía trascendental kantiana en filosofía analítica”¹². Por tanto, según dichos autores:

Muchos de los rasgos distintivos de la filosofía analítica están ya presentes en Peirce, y muchos de los problemas que más le preocuparon están ahora en el centro del debate filosófico contemporáneo. En este sentido, la mejor aproximación a Peirce sería suponer que él buscaba hacer algo parecido a la filosofía analítica, pues tanto sus objetivos como sus ideas básicas tienen mucho en común con este último¹³.

No obstante, Nubiola señala que cuando se confrontan específicamente los textos de Wittgenstein no aparece una sola mención a Peirce; tanto en el prólogo del *Tractatus* como en el de las *Investigaciones Filosóficas* se excusa Wittgenstein por no emplear el aparato bibliográfico tradicional en las exposiciones de carácter académico. Sin embargo, Nubiola documenta cómo, en los últimos años de su vida, Wittgenstein tuvo un interés considerable en el pragmatismo de James y lo leyó personalmente: “Con alguna frecuencia Wittgenstein se refiere a él en sus lecciones para ilustrar las confusiones en filosofía de la mente, y al parecer durante cierto tiempo los *Principios de Psicología* de James era el único libro que conservaba en su habitación”¹⁴.

En consecuencia, el ensayo de Nubiola busca rastrear hasta qué punto el pensamiento de Peirce influyó, a través de diversos autores que estuvieron en contacto con ambos filósofos, en el pensamiento de Wittgenstein. Así, en primer lugar, Nubiola hace una presentación de los datos disponibles acerca de la recepción de Peirce en la filosofía británica de las tres primeras décadas del siglo, prestando especial atención a Lady Welby, Charles Ogden y Bertrand Russell. En segundo lugar, da noticia con cierto detalle de la mediación de Ramsey como vínculo de

¹² *Ibidem*.

¹³ *Ibidem*.

¹⁴ *Ibid.*, p. 282.

unión entre Peirce y el segundo Wittgenstein; finalmente, a modo de conclusión, hace una breve estimación general de la literatura académica en torno a las relaciones entre los dos pensadores.

A continuación, se reseña críticamente dicha investigación, ya que se considera indispensable para poder diferenciar el punto de vista de este artículo de las opiniones de otros autores. Posteriormente, se examinará críticamente a otros autores a partir de una lectura directa de sus textos.

Puesto que la influencia de Russell sobre Wittgenstein está claramente establecida, para Nubiola es importante llamar la atención sobre la influencia de Peirce sobre Russell, vía los trabajos de Lady Welby y de F. C. S. Schiller, como antecedente de la evolución filosófica de Russell hacia el problema de la relación entre el lenguaje y los hechos, ya que “[Russell] en los *Principia Mathematica* había considerado el lenguaje como algo transparente y por tanto que podía usarse sin prestarle atención”¹⁵.

Se sabe que Russell reconoció en los *Principia Mathematica* la importancia de la lógica de relaciones de Peirce. No obstante, según Nubiola, en la primera década del siglo XX Russell no manifiesta el menor interés por la semiótica peirceana, a pesar de los intentos por conectar ambos autores de Lady Welby, “que creía encontrar en los trabajos de Russell y Peirce una confirmación de sus teorías semánticas”¹⁶. Sin embargo, con el paso de los años, Russell escribe:

Se ha hecho más por el avance [de la lógica formal] en cada una de las décadas que han seguido a 1850 de lo que se hizo en el período que va de Aristóteles a Leibniz (...) Se ha inventado una nueva rama de la lógica, la llamada Lógica de Relaciones (esta se debe en lo fundamental a C.S. Peirce), para tratar temas que estaban más allá del poder de la lógica aun cuando formaban parte del contenido fundamental de las matemáticas¹⁷.

En consecuencia, no es muy arriesgado argumentar que si Peirce influyó indirectamente en Russell y éste a su vez influyó en Wittgenstein, entonces, Peirce pudo haber influido en el pensamiento de Wittgenstein.

¹⁵ *Ibid.*, p. 294.

¹⁶ *Ibid.*, p. 285.

¹⁷ RUSSELL, B. “Los metafísicos y las matemáticas”. En: NEWMAN, J. R. (Ed.). *Sigma: el mundo de las matemáticas*. Vol. 4. Barcelona: Editorial Grijalbo, 1979. p. 369.

No obstante, el anterior argumento es aún débil. Para fortalecerlo hay que encontrar más conexiones. En su apoyo, Nubiola aporta la bien documentada relación de Ramsey con Wittgenstein, influencia que es ampliamente reconocida y agradecida por éste último en el prólogo de las *Investigaciones filosóficas*: “Advertir estos errores me ha ayudado – en un grado que apenas yo mismo puedo apreciar – la crítica que mis ideas han encontrado en Frank Ramsey, con quien las he discutido durante los dos últimos años de su vida en innumerables conversaciones”¹⁸.

A su vez, la conexión de Ramsey con las ideas de Peirce resulta más directa: “Ramsey habría tenido conocimiento de [una parte de] la lógica de Peirce a través de Russell, de Schroder y quizá de la extensa presentación que desarrolló C. I. Lewis en *A Survey of Symbolic Logic* [de] (1918)”¹⁹. Según Nubiola: “Thayer sugiere que Ramsey quizá pudo tener noticia por Russell del interés de Peirce por la teoría de los signos y el significado²⁰, pero seguramente fue Ogden, con el que a los 18 años de edad había preparado la traducción del *Tractatus*, quien le dio noticia más amplia de los escritos de Peirce”²¹.

A continuación Nubiola reseña los testimonios de Ramsey que acreditan su relación con los textos peirceanos:

En la amplia y valiosa reseña que Ramsey publica del *Tractatus* en la revista *Mind*, la única persona citada además de Russell y el propio Wittgenstein, es precisamente Peirce. Ramsey advertía²² que el uso de “*proposition*” en el *Tractatus* era ambiguo y que se habría podido evitar esa ambigüedad si Wittgenstein hubiera introducido la distinción usada por Peirce entre tipo (*type*) e instancia o réplica (*token*).

Sabemos que Ramsey ya había escrito esta reseña antes de visitar a Wittgenstein en Puchberg, pero no se sabe si hablaron de esta observación en las visitas de 1923 y 1924²³.

¹⁸ NUBIOLA, Op. cit., p. 286.

¹⁹ *Ibidem*. Se coloca entre corchetes “[una parte de] la lógica de Peirce” porque no incluye su – como él mismo la llama – *chef d’oeuvre* (obra maestra) en lógica: los gráficos existenciales.

²⁰ THAYER, H. citado por NUBIOLA, Op. cit., p. 286.

²¹ *Ibidem*.

²² Cfr. RAMSEY, F. “Review of *Tractatus Logico-Philosophicus*”. En: *Mind*, 1923. vol. 32, p. 464-478. Citado por NUBIOLA, Op. cit., p. 286.

²³ *Ibid.*, p. 287.

Posteriormente, “en la recensión que Ramsey publica al año siguiente en *Mind* de *The Meaning of Meaning* de Ogden y Richards, destaca que ‘merece especial atención el excelente apéndice sobre Peirce’”²⁴. Según Nubiola, también Ramsey reconoce la influencia de Peirce en aspectos cruciales de su propio pensamiento:

En “*Truth and Probability*”, leído en Cambridge y compilado póstumamente en sus *Philosophical Papers*, Ramsey cita en tres ocasiones a Peirce, y basa explícitamente en los textos de Peirce sus últimos párrafos sobre la justificación pragmática de los hábitos intelectuales de la inferencia, la observación y la memoria²⁵.

Ramsey sostiene que la inducción es un hábito y que tal hábito no requiere una justificación lógica, pues no puede darse ninguna justificación que no emplee la inducción²⁶.

Para otro autor, Sahlin, este trabajo supondrá en la corta vida intelectual de Ramsey el “comienzo de algo nuevo” pues está “imbuido del pragmatismo de Peirce”²⁷. Ramsey cita los textos de Peirce por la antología *Chance, Love and Logic*. Sin embargo,

En “*Facts and Propositions*” Ramsey no cita a Peirce, pero en los párrafos finales después de destacar su importante deuda respecto de Wittgenstein “de quien procede mi concepción de la lógica” añade:

“Debo a él todo lo que he dicho, excepto aquellas partes que tienen una tendencia pragmatista, que me parece son necesarias para llenar un hueco en su sistema”²⁸.

Supuestamente el reconocimiento de un vacío en el *Tractatus* que pudiera ser cubierto por el pragmatismo refuerza la hipótesis — aceptada por Nubiola — que Ramsey en sus conversaciones con Wittgenstein tratara de darle a conocer tanto el problema que dicho vacío implica, como la solución pragmatista que él había esbozado. No obstante, es preciso advertir, como señala Nubiola, que en este pasaje, de modo totalmente sorprendente,

²⁴ *Ibidem*.

²⁵ RAMSEY, F. *Philosophical Papers*. MELLOR, D. H. (Ed). Cambridge: Cambridge University Press, 1990. Citado por NUBIOLA, *Ibidem*.

²⁶ NUBIOLA, *Ibidem*.

²⁷ SAHLIN, N. *The Philosophy of F. P. Ramsey*. Cambridge: Cambridge University Press, 1990. Citado por NUBIOLA, *Ibidem*.

²⁸ NUBIOLA, *Ibidem*.

Ramsey atribuye su pragmatismo a Russell y no a Peirce:

“Mi pragmatismo procede de Russell, y por supuesto, resulta muy vago y no desarrollado- la esencia del pragmatismo me parece la siguiente; que el significado de una oración ha de ser definido por referencia a las acciones a que conduciría su aserción, o más vagamente todavía, por sus causas y efectos posibles”.

La sorprendente mención de Russell como origen del pragmatismo de Ramsey resulta sin duda algo desconcertante, ya que Russell no es un pragmatista ni nunca ha sido así considerado. Russell conocía los trabajos de James y algunos de Peirce, pero en sus publicaciones no había defendido nunca una posición pragmatista²⁹.

En realidad, lo que sucede es que la afirmación de Sahlin, en el sentido de que Ramsey “estaba imbuido del pragmatismo de Peirce”, sencillamente no parece ser correcta. En efecto, observemos la cita con la que prácticamente termina Ramsey su crítica a Wittgenstein y que, en nuestra opinión, marca claramente su desconocimiento del más profundo Peirce:

El peligro principal para nuestra filosofía –aparte de la pereza y de la confusión– es el escolasticismo, cuya esencia es tratar lo que es vago como si fuera preciso y tratar de meterlo en una categoría lógica exacta. Una muestra típica de escolasticismo es la tesis de Wittgenstein de que nuestras proposiciones de cada día están completamente en orden y que es imposible pensar ilógicamente³⁰.

Curiosamente lo que Ramsey –supuestamente “imbuido del pragmatismo de Peirce” – critica en Wittgenstein, a saber: el escolasticismo “cuya esencia es tratar lo que es vago como si fuera preciso y tratar de meterlo en una categoría lógica exacta”, es precisamente uno de los más importantes sostenes (eso sí, después de ser filtrado por sus herramientas semiótico-modales) de la arquitectónica peirceana, a saber: el realismo escolástico de los universales. En el apartado siguiente se dirá más sobre ello.

²⁹ *Ibidem*.

³⁰ *Ibidem*.

En síntesis, la evaluación de la investigación académica de la relación entre Wittgenstein y Peirce, según Nubiola, se deja resumir en los siguientes seis puntos:

1. La mención más antigua que se ha encontrado de la posible influencia de Peirce en la conformación del pensamiento del segundo Wittgenstein a través de Ramsey, se remonta a 1961. Se trata de una conjetura del australiano Gasking en una conferencia en la Universidad de Illinois en la primavera de aquel año. Albert Mullin, al registrar esto, daba noticia como contraste de una comunicación personal de Bertrand Russell en la que éste decía dudar que Peirce pudiera haber tenido alguna influencia sobre Wittgenstein. Por su parte Mullin venía a concluir no sólo que era improbable que uno hubiera influido en el otro a causa de sus bien diferentes estilos filosóficos, sino que sus pensamientos son más complementarios que similares³¹.

2. En años sucesivos, autores como Thayer, Hardwick, y Deledalle, defendieron la tesis de que Ramsey habría dado a conocer a Wittgenstein las ideas de Peirce. Esto haría más comprensible – aunque quizá tampoco los explicaría del todo – los aspectos claramente pragmáticos del segundo Wittgenstein. En este sentido Hardwick sostenía en 1979 que la importancia de la influencia de Peirce en Wittgenstein a través de Ramsey no había sido plenamente explorada y puso especial énfasis en que “lo que se necesita es un cuidadoso estudio de los temas comunes a Peirce y a Wittgenstein. Un estudio así proporcionaría un contexto de interpretación que añadiría una dimensión importante para la comprensión del trabajo del segundo Wittgenstein”³². Según Nubiola, en los últimos tiempos se ha realizado algún trabajo en esta dirección, a pesar de la dificultad que supone el que Peirce y Wittgenstein sean el centro de atención de dos comunidades de investigadores muy diferenciadas y con apenas conexión entre sí.

3. Joseph Ransdell destacó, frente a la llamativa disparidad de ambos filósofos respecto de sus actitudes ante la ciencia, la tradición filosófica o el carácter sistemático del pensamiento, el acuerdo fundamental entre sus filosofías, que podrían caracterizarse como dos versiones

³¹ Cfr. *Ibid.*, p. 290.

³² HARDWICK, C. S. Op. cit. Citado por NUBIOLA, Op. cit., p. 290. De Hecho, si este trabajo no se considera fallido, lo que intentamos en parte es responder a la petición de Hardwick, aunque motivados no tanto por la “comprensión del trabajo del segundo Wittgenstein” sino, más bien para su re-interpretación, pero esta vez, a la luz de una mejor comprensión del pragmatismo peirceano.

—la de Wittgenstein más escéptica y la de Peirce más optimista— de una común concepción de la razón y del lenguaje humano, entendidos esencialmente no como propiedad privada de los individuos, sino más bien como pensamiento ejercido comunitariamente³³.

4. Gullvág en 1981 dio gran realce a las semejanzas entre Wittgenstein y Peirce para llamar la atención sobre “la posibilidad de que los escritos de Peirce influyeran indirectamente en Wittgenstein a través de Ramsey, y de que, después de la muerte de Ramsey, pudieran haberle influido a él directamente”³⁴.

5. Renford Bambrough llamó la atención sobre las amplias coincidencias entre ambos pensadores. Cuando Wittgenstein recuerda en el párrafo 81 de las *Investigaciones Filosóficas* que “Ramsey insistía un día en que la lógica era una ciencia normativa”, el estudioso de Peirce no puede menos que pensar en que Ramsey estaba usando en aquella ocasión el eslogan de Peirce, aunque la expresión no fuera original de Peirce y aunque la idea se encuentre en otros autores que no emplean esa expresión. Esta impresión se refuerza al advertir en las *Investigaciones* numerosos ecos de ideas, de expresiones, de analogías y comparaciones que ahora son bien conocidas gracias a los ocho volúmenes de los *Collected Papers* de Peirce³⁵.

6. Christopher Hookway, en 1985, da luz sobre un tema capital para la comprensión de la filosofía peirceana: el análisis de la vaguedad y la indeterminación. Hookway —que había sido precedido en esta línea entre otros por los trabajos de Rorty (1961) y Fairbanks (1964)— mostraba que Peirce, Ramsey y el segundo Wittgenstein no sólo coinciden en afirmar que la vaguedad y la indeterminación del sentido de los predicados es benigna y tolerable, sino que incluso los tres vienen a encontrarse en la defensa de que la vaguedad “es una virtud, algo en cuya ausencia seríamos incapaces de decir, pensar, o hacer lo que queremos”³⁶.

³³ NUBIOLA, *Ibidem*.

³⁴ *Ibid.*, p. 291.

³⁵ *Ibidem*.

³⁶ *Ibidem*. Nótese que en este punto los tres autores y Bell hacen una afirmación que contradice la anterior crítica de Ramsey a Wittgenstein por querer “(...) tratar lo que es vago como si fuera preciso (...)”, afirmación que parece apoyar la crítica que más arriba se le hace a Sahlin cuando le atribuía a Ramsey “estar imbuido” de la filosofía de Peirce. Sin embargo, cabe también la posibilidad de que Ramsey haya modificado con el tiempo su opinión acerca de la vaguedad.

En resumen, está sin aclarar, a falta de una prueba concluyente, una conexión directa entre Peirce y Wittgenstein. Según Nubiola:

Hay un cierto número de conexiones a partir de autores que en diversos momentos tuvieron algún conocimiento del pensamiento peirceano y estuvieron en contacto con Wittgenstein, pero en especial falta el eslabón perdido que estaría constituido por el conocimiento del contenido de las conversaciones de Ramsey con Wittgenstein en los dos últimos años de vida de aquel.

No obstante:

No es fácil decir qué debe Wittgenstein a Peirce. Aun así resulta plausible afirmar que la vena pragmática del segundo Wittgenstein es de naturaleza peirceana desde el momento en que Wittgenstein leyó a James y fue ampliamente influenciado por Ramsey, quienes a su vez tuvieron contactos muy directos con Peirce³⁷.

Una vez terminada la reseña del ensayo de Nubiola, a continuación se presenta el análisis de otros autores iniciando con Rossella Fabbrichesi. Para ella "Peirce y Wittgenstein son preferiblemente cercanos por los aspectos pragmáticos de su filosofía o, dicho muy sintéticamente, por la teoría según la cual el significado de una palabra equivale a su uso en la práctica de la vida cotidiana, a los hábitos que ella produce"³⁸.

Además de éste:

Existen otros elementos a partir de los cuales se puede establecer una provechosa conexión entre los dos filósofos; su común planteamiento anticartesiano en filosofía, antipsicologista en lógica, el común acercamiento fenomenológico de los datos de la experiencia, y la final aspiración "metafísica" de su filosofía³⁹.

³⁷ *Ibidem*.

³⁸ FABBRICHESI, R. "L'immagine logica in Peirce e in Wittgenstein". En: BONFANTINI, M. A. & MARTONE, A. (Eds.). *Peirce in Italia*. (p. 393-398). Napoles: Liguori, 1993. p. 393.

³⁹ *Ibidem*.

Fabbrichesi, sin embargo, cree que existe:

Un plano más profundo que una posible conexión entre estos dos pensadores y al mismo tiempo más significativo, si sólo se presta atención a las numerosas páginas que ellos han dedicado al problema de la búsqueda lógica, o mejor aún, a la explicación teórica de los fundamentos de la lógica. Peirce hubiera podido suscribir, creo, la famosa afirmación de Wittgenstein: —Sí, mi trabajo se ha extendido desde los fundamentos de la lógica a la esencia del mundo—⁴⁰.

Según Fabbrichesi:

Más allá de las innegables contribuciones que ellos han aportado a la historia del pensamiento formal, la anterior afirmación se puede considerar como el punto más destacado de sus investigaciones. Su actividad lógica es particularmente poco ortodoxa: sus más importantes descubrimientos de tipo técnico se encuentran frecuentemente incluidos en un cuadro de reflexión teórica general y de pretensión filosófica, las cuales difícilmente pueden ser distanciadas, no obstante los tentativos hechos en este sentido de parte de intérpretes como Russell y Schroder⁴¹.

¿Cuál es, en efecto, el objeto de investigación lógica en Peirce y Wittgenstein? Para Fabbrichesi podría ser expresado en forma muy general del siguiente modo:

Peirce y Wittgenstein están en la búsqueda del sentido de la representación; en términos de Wittgenstein, del sentido de la proposición o representación lógica; en términos de Peirce, del sentido de la significación, de la transmisión sígnica. Podríamos preguntar en modo un poco tosco, pero sintético: ¿Qué es, o mejor, qué hace una representación cuando muestra el mundo al pensamiento, concentrándolo en una imagen lógica? O también decir, con la colorida exclamación de Wittgenstein: —¿Pero qué diablos es este lugar lógico?—⁴².

⁴⁰ *Ibidem*.

⁴¹ *Ibidem*. En rigor, la comparación que hace Fabbrichesi es equivocada e injusta porque “igual” los reconocidos y significativos descubrimientos lógicos de Peirce, con unos supuestos (?) descubrimientos lógicos de Wittgenstein, cuando se sabe que este último no realizó ningún descubrimiento técnico significativo a parte de las tablas de verdad.

⁴² *Ibid.*, p. 394.

Según la autora:

Peirce y Wittgenstein responden de modo similar a esta pregunta y llegan a una misma conclusión, que podría ser articulada en los siguientes puntos:

1. El lugar lógico es ante todo un lugar icónico. Una imagen, un *Bild*, en los términos de Wittgenstein; un icono, o sea una *Primeridad* en los términos de Peirce. Un lugar de reduplicación, entonces, de semejanza.
2. Al ser un lugar de semejanza éste se entiende como un espacio de relaciones internas, o sea de relaciones no relativas. Veremos qué entienden los dos autores con esta expresión que los une también en un nivel terminológico.
3. El lugar lógico es forma de estas relaciones, o sea forma de la posibilidad. Lógica y posibilidad están estrechamente unidas, y los dos autores centran sobre esto su atención.
4. En fin, el lugar lógico se configura como un espacio limitado e invadido por la ética, desde lo místico, dice Wittgenstein; desde la *Primeridad* y la *Segundidad*, dice Peirce. No olvidemos que para los dos lógicos, ética y estética sostienen un profundo lazo y son consideradas relacionalmente⁴³.

Brevemente, Fabbrichesi anota sobre los primeros dos puntos:

Que la gran intuición de estos dos pensadores coincide en que la lógica revela en su originaria constitución, un alma icónica: o como dice Peirce, los primeros signos son iconos. Wittgenstein afirma, como es de notar, que la lógica no es una doctrina, sino una imagen especular del mundo. Y el Peirce de los Gráficos Existenciales tenía una idea análoga: “toda afirmación contiene un icono”⁴⁴, enfatizando de hecho la *l'Abbildungstheorie* de Wittgenstein y un icono equivale a cualquier cosa “que pueda ser sustituida por cualquier otra que se le asemeje”⁴⁵.

⁴³ *Ibid.*, p. 394-395. Cursivas añadidas.

⁴⁴ PEIRCE, C. S. CP 2.280 citado por FABBRICHESI. *Ibid.*, p. 395.

⁴⁵ FABBRICHESI. *Ibidem*.

O también, en un contrapunto entre citas de Wittgenstein y Peirce, “‘la figura lógica de los hechos es el pensamiento’, dice la tercera proposición del *Tractatus*. [O] ‘la belleza de los gráficos lógicos está en que sean verdaderamente Icónicos’, afirma Peirce”⁴⁶.

Lo lógico es entonces, un lugar de semejanza, de reproducción ‘en imagen’ del mundo, de duplicación. Ahora, ¿cómo se da una representación lógica? “En imagen y representación – escribe Wittgenstein – algo debe ser idéntico, con el fin de que ésta sea una imagen de éste. Lo que toda imagen, de cualquier forma que sea, debe tener en común con la realidad para poder representar – correcta o falsamente – es la forma lógica, o sea la forma de la realidad”⁴⁷.

Por tanto:

Esta identidad que debe reproducirse como tal en el reflejo entre mundo y realidad para garantizar la significación, es entonces en Wittgenstein, la forma de la representación, mientras en Peirce podría identificarse, con el *ground*, que es precisamente medida de la similaridad. “No podemos comprender un acuerdo entre dos cosas, sino como acuerdo en cualquier aspecto, y esto es una abstracción pura o *ground*”⁴⁸.

Pero en general, el icono de Peirce contiene caracteres casi⁴⁹ idénticos al *Bild* de Wittgenstein: “‘Este representa la lógica porque está gobernada por la misma ley’. No tiene distinciones entre sí y el objeto, porque se identifica con la cosa misma; es, sí una imagen, pero estrictamente hablando puede ser sólo una idea o aún mejor, una posibilidad, una categoría primaria”⁵⁰.

Y continúa:

Identificando icono y *Primeridad*, Peirce debilita la base de la teoría “ingenua” de la representación, ya que destruye las bases lógicas: el icono no se adapta a una realidad previamente constituida, sino que establece la primera

⁴⁶ *Ibidem*.

⁴⁷ *Ibid.*, p. 396.

⁴⁸ *Ibidem*.

⁴⁹ Se dice *casi* porque Fabbrichesi no menciona la fundamental triadicidad explícita de la concepción peirceana del signo.

⁵⁰ *Ibidem*.

posibilidad del reconocimiento simbólico a través de una relación de similitud, en la cual los relatos se caracterizan en la recíproca identificación. El icono sustituye y en conjunto constituye el propio objeto, desarrollando así la misma función que Peirce asigna a toda representación. En efecto, “las cosas reales son de naturaleza cognoscitiva y, por ello, significativa”⁵¹, el primer iconismo como ya entreveía Wittgenstein, está entre forma lógica y forma de la realidad. Ahora, esto no significa afirmar vanamente que las proposiciones lógicas deban asemejarse a la realidad tal como es; indica más bien la concepción común por parte de los dos autores de cómo el problema de la lógica es el problema de la semejanza, o sea del tautón entre ser y pensamiento, que se refleja en cuanto la forma entre ellos es idéntica y se fundan recíprocamente. La lógica colma el mundo, porque más allá de las señales con las que lo indicamos, no es para el pensamiento un mundo⁵².

Según Fabbrichesi:

Al observar entonces, más que la identidad entre signo y designado, sería mejor hablar de semejanza, entendiéndola como relación interna o no relativa. Esta curiosa expresión indica en los dos autores, la particular relación de retroalimentación gracias a la cual la imagen envía nuevamente al propio objeto. Ésta establece no una diferencia real entre dos hechos opuestos entre ellos, sino “una comunidad de caracteres entre dos elementos que no se distinguen aún como tales”⁵³.

Obsérvese que para la autora entre signo (icono) y objeto, o dicho en otras palabras, entre lenguaje y mundo hay una relación interna, vale decir, una relación de similitud: “abstracta de lo concreto que implica la posibilidad de otro”⁵⁴.

Pero, de nuevo, se echa de menos la referencia de Fabbrichesi a la terceridad en el signo peirceano en contraste con la dualidad wittgensteniana:

⁵¹ PEIRCE, C. S. CP 5.320 citado por FABBRICHESI, *Ibidem*.

⁵² FABBRICHESI. *Ibidem*.

⁵³ *Ibid.*, p. 397.

⁵⁴ *Ibidem*.

El objeto de la imagen lógica en Peirce y Wittgenstein, es su mismo (*self*) icónico: y es desde este *ground* común, en que cualquier cosa se reconoce como similar, que se abre la posibilidad de la relación de significación, de la lógica, del sentido. Ya que el lenguaje se da al mundo en relaciones internas, aquél o estas relaciones determinan la posibilidad de los hechos.

(...) Como dice Wittgenstein, el lugar lógico entendido como un nexo de relaciones, determina la posibilidad lógica de los hechos, Peirce diría: objeto del icono es la forma de la relación representada, o sea la forma de una posibilidad lógica. El icono es un posible que representa un posible.

Lógica, imagen y *posibilidad* están estrechamente unidas⁵⁵.

Con las cursivas se busca llamar la atención sobre el hecho que Fabbrichesi ha señalado —aunque vagamente—⁵⁶ en su artículo, un aspecto que se encuentra en línea con la posición de este trabajo, a saber, el sentido del mundo para Peirce no puede ser esclarecido sin una aproximación semiótico-modal; aún así, aquí se considera que la autora se equivoca al afirmar que:

El sentido del mundo tanto para Peirce como para Wittgenstein, consiste en la posibilidad de su representación icónica; Wittgenstein encuentra el cómo del mundo en el espejo del lenguaje, Peirce lo encuentra en los signos. Pero qué es mundo —la cualidad de su ser— es para ambos inefable e indescriptible: se puede mostrar icónicamente, mas no se puede decir. Dicho en otras palabras: Peirce y Wittgenstein miran constantemente, por ser lógicos, además de los confines de lo lógico, las distancias y límites, confrontándolos con la falta de palabras adecuadas para decirlo⁵⁷.

Más abajo se intentará mostrar que en Peirce, si es que encontramos las palabras adecuadas para decirlo, no hay signo con sentido de lo absolutamente incognoscible.

⁵⁵ *Ibid.*, p. 398. Cursivas añadidas.

⁵⁶ Se dice *vagamente* porque en ninguna parte de su artículo la autora hace alguna mención explícita de la modalidad. Y sin embargo, se encuentra entre los pocos autores que han visto —así sea de manera vaga— aspectos claves de la modalidad peirceana al interior de su semiótica universal.

⁵⁷ *Ibidem*.

En un artículo más reciente John Upper⁵⁸ llama la atención sobre la “demanda histórica” de Von Wright y Hanfling de que las ideas de Wittgenstein, por ser muy diferentes a cualquier cosa vista antes, no podrían resistir comparación alguna con autores precedentes y, sobre la “demanda conceptual”, de Davidson, de que no es posible comparar pensadores con ideas tan disímiles. No obstante, Upper considera que estas demandas no son tan válidas ya que, según él, las ideas de Wittgenstein no son tan distintas a las de Peirce.

De todas maneras, para Upper, “el problema no es si Wittgenstein y Peirce compartieron muchas de estas mismas creencias, sino sí esas creencias compartidas son triviales o importantes”⁵⁹. El punto de esta observación es que señala “que hay una tentación real en exagerar las similitudes entre dos pensadores enfocando todo el esfuerzo [de comparación] en un sólo aspecto de su trabajo”⁶⁰.

Para evitar los errores de comparación a los que alude Upper, él ha escogido concentrarse en los aspectos del trabajo de Wittgenstein que la comunidad filosófica parece aceptar como contribuciones importantes y distintivas a la filosofía. En ese sentido, Peirce y Wittgenstein tienen concepciones similares de “significado” cuando enfatizan la relación entre significar y acción práctica:

Las similitudes se extienden a otras demandas que son fundamentales a sus posiciones: los dos reconocieron la necesidad de clarificar los conceptos, los dos reconocieron la importancia de la vaguedad en el idioma ordinario, los dos rechazaron la duda cartesiana y vieron el escepticismo como un “tigre de papel”, y los dos insistieron en el criterio público de *meaningfulness*⁶¹.

Para otro autor, Jaakko Hintikka, más cercano al verdadero pensamiento peirceano, dice que:

A semejanza de Peirce, Wittgenstein creyó que el uso del lenguaje constituye su significado. Y a semejanza de Peirce, Wittgenstein no se refiere mediante “uso” a un mero

⁵⁸ Cfr. UPPER, J. On the very idea of comparing Wittgenstein and Peirce. [En línea]: <http://www.cspeirce.com/menu/library/aboutcsp/upper/scholar.htm>

⁵⁹ *Ibidem*.

⁶⁰ *Ibidem*.

⁶¹ *Ibidem*.

“uso” verbal, esto es, a un “juego” cuyas jugadas son actos lingüísticos. Ambos enfatizan el uso de manera pragmática en el sentido de la utilización del lenguaje. (Habla a favor de esto, por ejemplo, la manera en la que Wittgenstein compara el lenguaje con una caja de herramientas). *Pero aquí terminan las semejanzas. Como universalista y en completo contraste con Peirce, Wittgenstein no podía decir de manera oficial nada acerca de los juegos de lenguaje en general o incluso desarrollar una teoría real de algún tipo particular de juego de lenguaje*⁶².

Para entender en qué sentido el autor usa el término ‘universalista’ hay que decir que Hintikka distingue entre dos posiciones o interpretaciones diferentes de la relación entre el lenguaje y el mundo, a saber, el lenguaje como el medio universal, o posición universalista, a diferencia del lenguaje como cálculo, o posición basada en la teoría de modelos. Para Hintikka,

“la posición universalista fue adoptada por Frege, el joven Russell, y el Wittgenstein de mediana edad, el Círculo de Viena de comienzos de los años treinta y, en cierto sentido Quine, Heidegger, y las tradiciones hermenéuticas y deconstructivistas, por el contrario, la posición basada en la teoría de modelos incluye personas tales como Boole, Peirce, Schroder, Lowenheim, Godel, Tarski, [entre otros]”⁶³.

Pero, ¿por qué es pertinente establecer dicha diferencia? Según Hintikka, “una consecuencia particular de la posición universalista es que nuestro lenguaje y su lógica no pueden auto-referirse como un todo, ni discutirse en su totalidad en un metalenguaje separado (a excepción de sus características puramente formales, por supuesto)”⁶⁴.

Por el contrario, para un pensador fiel a la tradición de lo que posteriormente se conocería como teoría de modelos, por ejemplo Peirce, la consecuencia anterior sería la contraria. En efecto, según Hintikka, un ejemplo de ello lo aporta el propio Peirce en un pasaje de su teoría de los gráficos existenciales:

Ahora pasaré a otra esfera de los grafos gamma de la cual es imposible prescindir. En efecto, es necesario que seamos

⁶² HINTIKKA, J. El viaje filosófico más largo: de Aristóteles a Virginia Woolf. Barcelona: Editorial Gedisa, 1998. p. 239-240. Cursivas añadidas

⁶³ *Ibid.*, p. 218-219.

⁶⁴ *Ibid.*, p. 226.

capaces de razonar en grafos acerca de grafos. La razón es que un razonamiento acerca de grafos consistirá de manera *necesaria* en mostrar que algo es verdadero de todo posible grafo correspondiente a una cierta descripción general. Pero no podemos trazar todo *posible* grafo correspondiente a cualquier descripción general y, por lo tanto, si hemos de razonar en grafos tenemos que tener un grafo que es una descripción general del tipo de grafo con el cual hay que relacionar el razonamiento⁶⁵.

Desde luego, como dice Hintikka, “si alguien toma con seriedad las modalidades y la lógica modal, tiene que aprobar alguna versión de la concepción del lenguaje como cálculo”⁶⁶ y continúa diciendo:

1. Al menos en su pensamiento maduro, Peirce comprendió las modalidades de manera realista en un verdadero sentido metafísico, sin tratar de justificarlas a través de rodeos, p. ej., de una manera epistémica.
2. En su obra real sobre lógica, Peirce desarrolló formas de estudiar lógicas modales. Por lo general, en su teoría lógica Peirce empleó de manera libre ideas que presuponen una multiplicidad de mundos posibles (u otros *possibilia*).
3. Peirce influyó sobre el desarrollo subsiguiente de los sistemas explícitos de lógica modal. En particular, fue muy admirado por C. I. Lewis.
4. La cuestión profunda aquí desde un punto de vista filosófico es que Peirce mostró una conciencia, mucho más penetrante que los miembros de la tradición universalista, de la distinción decisiva entre verdad (verdad *simpliciter*) y el mal llamado concepto de “verdad lógica”⁶⁷.

La cita anterior es decisiva para la evaluación que se hace en este artículo de la investigación académica de la relación entre Peirce y Wittgenstein porque apoya una posición tercera entre los que consideran que los dos autores tienen ideas muy semejantes, entre ellos Upper, y otros como Davidson, que definitivamente niegan lo anterior. En efecto, la posición defendida en este trabajo es que las intuiciones correctas del pensamiento wittgensteniano, tal como las descubre Fabbrichesi, están encerradas en

⁶⁵ *Ibid.*, p. 226-227. Cursivas añadidas.

⁶⁶ *Ibid.*, p. 223.

⁶⁷ *Ibidem*.

una camisa de fuerza 'universalista', según indica Hintikka, y que con una relectura de Wittgenstein en clave peirceana, se pueden liberar de sus ataduras.

Así pues, se ha intentado mostrar cómo, en las aproximaciones al tema de la relación entre Peirce y Wittgenstein, con la excepción de Hintikka, no se puede ir más allá de los tópicos pragmáticos comunes tanto en Wittgenstein como al pragmatismo que James y Ramsey, entre otros, derivaron de Peirce.

Sin embargo, lo interesante del asunto es que todavía es posible contrastar a nuestros dos filósofos, pero ahora a la luz del pragmatismo que Peirce alcanzó a esbozar en sus líneas maestras durante las últimas dos décadas de su vida. Por eso se considera importante extraer, de la reseña de los trabajos anteriores, la conclusión de que no se debe pasar por alto el fundamental aspecto modal que una verdadera confrontación de la obra de Peirce, con cualquier otro autor, parece exigir.

REFERENCIAS

- DELEDALLE, G. (1996). *Leer a Peirce hoy*. Barcelona: Editorial Gedisa.
- FABRICCHESI, R. (1993). "L'immagine logica in Peirce e in Wittgenstein". En: BONFANTINI, M. A & MARTONE, A. Eds. *Peirce in Italia*. (p. 393-398). Napoles: Liguori.
- HINTIKKA, J. (1998). *El viaje filosófico más largo: de Aristóteles a Virginia Woolf*. Barcelona: Editorial Gedisa.
- NUBIOLA, J. (1996). "Scholarship on the relations between Ludwig Wittgenstein and Charles S. Peirce". En: ANGELELLI, I. & CERESO, M. (Eds.). *Proceedings of the III Symposium on History of Logic*. Berlin: Walter de Gruyter GmbH & Co.
- PEIRCE, C. S. (1931-1935). *Collected Papers of Charles Sanders Peirce*. 6 vols. HARTSHORNE, C. & WEISS, P. Eds. Cambridge, MA: Harvard University Press.
- _____. (1958). *Collected Papers of Charles Sanders Peirce*. Vols. 7-8. A. BURKS, W. (Ed.). Cambridge, MA: Harvard University Press.
- _____. (1968). *Escritos lógicos*. Madrid: Alianza Editorial.
- _____. *Un argumento olvidado en favor de la realidad de Dios*. [En línea]: <http://www.unav.es/gep./Peirce-esp.html>.
- PUTNAM, H. (1994). *Cómo renovar la filosofía*. Madrid: Ediciones Cátedra.

RAMSEY, F. (1923). "Review of *Tractatus Logico-Philosophicus*" En: *Mind*, vol.32, p. 464-478.

_____. (1990). *Philosophical Papers*. MELLOR, D. H. (Ed.). Cambridge: Cambridge University Press.

RUSSELL, B. (1979). "Los metafísicos y las matemáticas". En: NEWMAN, J. R. (Ed.). *Sigma: el mundo de las matemáticas*. Vol. 4. Barcelona: Editorial Grijalbo.

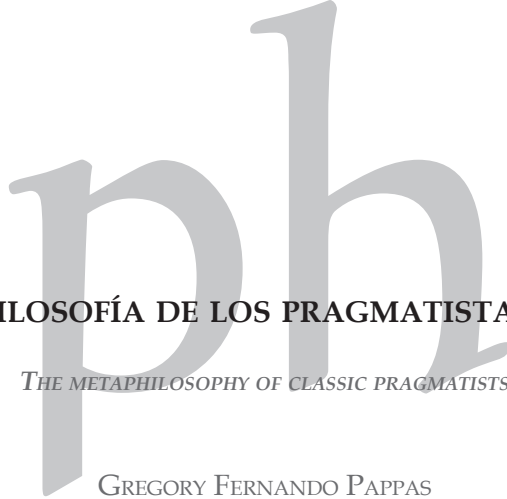
SAHLIN, N.E. (1990). *The Philosophy of FP Ramsey*. Cambridge: Cambridge University Press.

UPPER, J. B. On the very idea of comparing Wittgenstein and Peirce. [En línea]: <http://www.cspeirce.com/menu/library/aboutcsp/upper/scholar.htm>

WITTGENSTEIN, L. (1973). *Tractatus logico-philosophicus*. Madrid: Alianza Editorial.

_____. (1984). *Los cuadernos azul y marrón*. Madrid: Editorial Tecnos.

_____. (1988). *Investigaciones filosóficas*. Barcelona: Editorial Crítica.



LA METAFILOSOFÍA DE LOS PRAGMATISTAS CLÁSICOS

THE METAPHILOSOPHY OF CLASSIC PRAGMATISTS

GREGORY FERNANDO PAPPAS

Texas A & M University, Estados Unidos. g_pappas@neo.tamu.edu

RECIBIDO EL 15 DE SEPTIEMBRE DE 2010 Y APROBADO EL 30 DE SEPTIEMBRE DE 2010

RESUMEN ABSTRACT

En este ensayo argumento que la metafilosofía de los pragmatistas es la contribución más importante de estos filósofos a la historia de la filosofía y es también lo que los distingue de otros filósofos. Los filósofos clásicos americanos (Peirce, James, y Dewey) y los filósofos pragmatistas hispanos, Ortega y Gasset y Risieri Frondizi, propusieron que la filosofía debe de partir desde la experiencia, es decir, un punto de partida práctico. Después de explicar qué significa sostener que la experiencia es el punto de partida, examino las razones que tienen estos filósofos para sostener que la experiencia es el punto de partida apropiado si se quiere que la filosofía sea empírica y relevante.

In this paper I argue that the metaphilosophy of pragmatists is the most important contribution of these philosophers to the history of philosophy and is also what distinguishes them from other philosophers. The classic American philosophers (Peirce, James, and Dewey) and the Hispanic philosophers, Ortega y Gasset and Risieri Frondizi, proposed that philosophy must start from experience. After explaining what it means to take experience as the starting point, I examine the reasons that pragmatists philosophers have for holding the view that experience should be the starting point if philosophy is to be empirical and relevant.

PALABRAS CLAVE KEY WORDS

empírico, experiencia, metafilosofía, práctica, Pragmatismo, punto de partida.

empirical, experience, metaphilosophy, practice, Pragmatism, starting point.

“La referencia constante y permanente a la totalidad de la experiencia debe de ser una exigencia ineludible de todo filósofo que pretenda convertirse en conocimiento genuino”¹

“Empirical method is the only method which can do justice to this inclusive integrity of “experience”. It alone takes this integrated unity as the starting point for philosophic thought”²

I. LA METAFILOSOFÍA Y LA NATURALEZA DEL PRAGMATISMO

En la introducción a *Ensayos Sobre Metafilosofía*, Pablo Quintanilla observa que recientemente, en muchos círculos filosóficos, a la metafilosofía no se le ha dado la atención que se merece. Esto se debe quizás a que se ha asumido que es “más importante ir a los problemas mismos antes que quedarse en los prolegómenos” o porque se ha asumido “que los temas metafilosóficos centrales ya están suficientemente establecidos”. Ninguna de estas dos razones para desatender la metafilosofía las encuentro convincentes. La primera razón pierde de vista que como un filósofo trata de resolver los problemas que le conciernen, asume una metodología y por lo tanto asume una metafilosofía. No niego que es importante atender a los problemas específicos. Tampoco creo que sea necesario que un filósofo examine conscientemente su metodología cada vez que va a atender un problema. Sin embargo, la capacidad de una filosofía para resolver o disolver problemas depende de su metafilosofía, no en el sentido de una teoría explícitamente formulada por el filósofo sobre la filosofía, pero en la manera habitual e inconsciente de conducir su investigación filosófica. A pesar de la importancia de la metafilosofía, por alguna razón u otra, los filósofos no somos suficientemente críticos de nuestras maneras de investigar.

A muchos filósofos no les interesa la metafilosofía porque asumen, “que los temas metafilosóficos centrales ya están suficientemente establecidos”. Este es el peligro que corren los filósofos cuando hacen filosofía que está de moda, o que se consideran “*mainstream*”, en otras palabras filosofía dominante o normal durante una época en el mundo

¹ FRONDIZI, Risieri. *El Punto de Partida del Filósofo*. Buenos Aires: Editorial Losada, 1945. 2a. Ed., 1957. p. 122.

² DEWEY, John. *Experience and Nature. Later Works. Vol. 1.* Southern Illinois: University Press, 1925. p. 19.

académico. Pierden estos la capacidad de criticar su manera particular y “normal” de hacer filosofía. Este punto es importante, pues justifica el valor de contrastar la metafilosofía de los pragmatistas clásicos con las maneras más comunes y tradicionales de hacer filosofía. Esto es en parte la meta de este ensayo.

En la historia de la filosofía domina una manera general de hacer filosofía: desde el punto de partida teórico o desde lo que se revela a la conciencia de un sujeto como conocedor. Este punto de partida es tan común que lo comparten aún filosofías que se consideran antitéticas o enemigas. No estuviéramos conscientes de esa manera general de filosofar si no fuera por las críticas de algunos filósofos a este punto de partida teórico. Si no fuera por una minoría “pragmatista”, especialmente los filósofos clásicos americanos (Peirce, James, y Dewey) y los filósofos pragmatistas hispanos (Ortega y Gasset y Risieri Frondizi) quizás no nos hubiéramos dado cuenta de que hay una alternativa: partir desde la experiencia o el punto de partida práctico. Parte de la vigencia e importancia de los pragmatistas clásicos estriba en cuestionar lo que continúa tomándose como normal y por sentado, especialmente en este asunto sobre cuál debe de ser el punto de partida de la filosofía.

Considero que la metafilosofía de los pragmatistas clásicos es la contribución más importante de estos filósofos a la historia de la filosofía y es también lo que los distingue de otros filósofos. En otras palabras, quiero además defender una posición sobre la naturaleza del pragmatismo como filosofía.

Los estudios recientes sobre el pragmatismo están de acuerdo en ciertas cosas. En primer lugar, una narrativa sobre la historia y naturaleza del pragmatismo debe de reconocer las importantes diferencias entre los filósofos pragmáticos. En segundo lugar, nada podría ser más anti-pragmático que reclamar que los pragmatistas comparten una esencia. En tercer lugar, no existe una doctrina común llamada pragmatismo, aunque sí hay cierta superposición de ideas y métodos comunes que han sido pasados y modificados a través de las generaciones de esta familia histórica de filósofos. ¿Cuáles son, sin embargo, las ideas filosóficas fundamentales? Por ejemplo, ¿qué es lo que Dewey heredó de Peirce y James que provocó su propia reconstrucción del pragmatismo?

En todo relato histórico del pragmatismo cierta selectividad es inevitable y esto es un motivo de desacuerdos sobre la naturaleza del

pragmatismo. El que haya selectividad no significa que cualquier relato es tan bueno como cualquier otro. Propongo que cualquier relato sobre el pragmatismo debe estar basado y defendido en, al menos, dos bases: (a) evidencia textual que los pragmatistas comparten hasta cierto punto las ideas filosóficas presentadas como centrales, y que (b) estas ideas filosóficas sean las más dignas de resurrección o reconstrucción. (b) es importante y suele ser la base de la selectividad que mencionaba. Los desacuerdos sobre la naturaleza del pragmatismo no se pueden resolver con sólo apelar a los textos clásicos del movimiento. Porque aún cuando estamos de acuerdo acerca de lo que dice el texto, existe un desacuerdo sobre qué es lo más importante y qué podemos dejar de enfatizar o ignorar. Nuestra selectividad al leer está basada en las siguientes preguntas: ¿Cuál es la contribución más significativa del pragmatismo en la historia de la filosofía? ¿Cuál es la diferencia que hace la mayor diferencia?

Con estos antecedentes en mente, estoy listo para tomar una posición sobre la naturaleza del pragmatismo. El pragmatismo es una filosofía revolucionaria en la historia de la filosofía, ya que criticó el punto de partida moderno y en su lugar tomó la “experiencia” como el punto de partida de cualquier investigación filosófica. El legado más importante en el desarrollo y la historia del pragmatismo es meta-filosófico, es decir, es sobre cómo proceder en hacer filosofía. Esta interpretación del pragmatismo está (a) basada en una lectura generosa, honesta y comprensiva de los textos y (b) sostiene que esta interpretación hace del pragmatismo un movimiento radical (más allá de la filosofía moderna y aún de algunos “post-modernistas”), y un recurso promisorio o útil para la crítica y el adelanto de la filosofía actual.

No tengo aquí el espacio para defender (a), pero creo que la evidencia textual es abrumadora. Un texto es, por supuesto, sujeto a una pluralidad de interpretaciones razonables, pero a muchos lectores de los filósofos pragmáticos se les ha hecho bien difícil leer los textos sin poner al lado sus propios prejuicios teóricos. Cabe advertir que mi concepción del pragmatismo es diferente a la idea común que el pragmatismo es sólo una filosofía que nace de una teoría del significado o de la verdad. Se dice que la historia del pragmatismo de Peirce comienza con su “*Pragmatic Maxim*” entendida como un método para determinar el significado de conceptos, en su ensayo de 1878 *Cómo hacer nuestras ideas claras*. William James simplemente amplió la máxima pragmática de Pierce a la moral y la religión. Dewey, por el otro lado, simplemente

elaboró las implicaciones socio-políticas y educativas. Esta noción del pragmatismo nace de las interpretaciones inglesas del pragmatismo, y se sigue promulgando por los neo-pragmáticos o académicos que trabajan en la epistemología o filosofía del lenguaje.

El problema con esta historia del pragmatismo, centrada en la filosofía del lenguaje y la epistemología, es que es bastante estrecha. En esta historia, no hay ninguna mención del empirismo radical de James o de la insistencia de Dewey en tomar la “experiencia” como método. No niego que aquellos que defienden esta historia pueden proporcionar suficiente evidencia textual para apoyar su interpretación, pero ignoran lo que me parece más fundamental, radical, y prometedor de esta tradición filosófica. El pragmatismo entendido como una metafilosofía es más incluyente e interesante. La máxima pragmática sobre el significado de los conceptos es sólo una de las consecuencias de tomar la experiencia vivida como punto de partida. De acuerdo a este punto de partida, el significado de los conceptos (incluyendo el de la “verdad”) debe ser expresado en términos de experiencia y sus consecuencias. Dewey caracteriza el punto de partida, o el método, favorecido por el pragmatismo en términos de “el postulado de empirismo inmediato”: “si se quiere realmente describir cualquier cosa, la tarea consiste en describir como se experimenta su ser”³. Por lo tanto, el significado de cualquier concepto, incluyendo el de la “verdad”, está atado a su uso en la práctica, es decir, en la vida cotidiana.

No quiero negar las importantes diferencias entre James, Dewey y Peirce en cuanto a muchos asuntos filosóficos. Sin embargo, comparten una metafilosofía. Incluso su compromiso compartido con el falibilismo y el pluralismo, son una consecuencia de tomar la experiencia vivida como punto de partida, ya que una de las primeras cosas evidentes acerca de la experiencia es que es cambiante, plural, abierta hacia el futuro desconocido. La metafilosofía de los pragmatistas clásicos los distingue de muchos filósofos contemporáneos incluyendo los neo-pragmatistas (como Rorty, Putnam, y hasta Davidson, Goodman, etc.). A pesar de las continuidades históricas y similitudes entre estos últimos filósofos y por ejemplo Dewey, hay una diferencia fundamental: el punto de partida de estos es el lenguaje, por eso me parece apropiado el nombre “*linguistic pragmatism*” para distinguir estos pragmatistas de los clásicos. Todavía vivimos en una época dominada por el llamado “giro lingüístico”, pero

³ *Ibid.*, p. 158.

para Dewey y James el punto de partida de la filosofía debe de ser la experiencia y no el lenguaje.

La noción del pragmatismo como una filosofía que toma a la experiencia como punto de partida tiene ciertas ventajas. Nos da una base sólida para criticar los debates contemporáneos en la filosofía, que a menudo son "peleas familiares", porque comparten el mismo punto de partida. Más importante aún, esta noción del pragmatismo nos permite hoy criticar o mejorar la hipótesis o conclusiones de los pragmatistas clásicos, pues es posible que algunas conclusiones de estos filósofos necesiten ser re-examinadas utilizando su propia metafilosofía.

Hasta ahora he argumentado a favor de una metafilosofía basada en la experiencia para definir la naturaleza e importancia del Pragmatismo, pero queda mucho por aclarar. ¿Qué significa sostener que la experiencia es el punto de partida, y qué razones tienen estos filósofos para sostener que la experiencia es el punto de partida apropiado si se quiere ser empírico en la filosofía? A continuación contesto estas preguntas, sostengo que los pragmatistas compartieron: (1) una crítica similar a la filosofía moderna basada en su punto de partida, (2) una defensa de la experiencia como el punto de partida apropiado de la investigación filosófica.

II. LA CRÍTICA DEL PUNTO DE PARTIDA MODERNO

Para los pragmatistas clásicos hay que estudiar la historia de la filosofía para poder aprender de los errores del pasado, especialmente cuando se trata del punto de partida de la investigación filosófica. Este es el asunto más importante en la filosofía, ya que diferentes puntos de partida conducen a diferentes filosofías. Más importante que el contenido o conclusiones de una filosofía es el método de investigación. La historia del Pragmatismo se puede entender como la historia de una constante y desarrollada crítica de un punto de partida que sigue siendo favorecido por muchos filósofos. Peirce se refirió a este punto de partida como "cartesianismo" porque lo detectó en Descartes. En un artículo publicado en 1905 dice:

Filósofos de muy diversas tendencias proponen que la filosofía establezca su punto de partida desde uno u otro estado mental en que ningún hombre, y menos un

principiante en filosofía, se encuentra realmente. Uno propone que comience dudando de todo, y dice que hay una sola cosa que no puede dudarse, como si dudar fuera “tan fácil como mentir” [Shakespeare, Hamlet, acto 3, escena 2]. Otro propone que deberíamos comenzar observando “las primeras impresiones del sentido”, olvidando que nuestras percepciones mismas son el resultado de la elaboración cognitiva. Pero en verdad no hay sino un estado mental desde el que se puede “comenzar”, a saber, el preciso estado mental en el que uno en realidad se encuentra en el momento de “comenzar”⁴.

A veces no estoy seguro si el mismo Peirce entiende lo radical y significativo que fue su crítica a la tradición filosófica⁵. La crítica a la tradición de Peirce no es sólo a Descartes; pero a el punto de partida de la filosofía moderna. La mayoría de los problemas y doctrinas filosóficas que identifican a la filosofía moderna son el resultado de partir de un punto de partida común. William James y John Dewey describieron este punto de partida como el punto de partida subjetivo y contemplativo, pues se asume que somos sujetos o espectadores de un mundo antecedente y fijo que espera a ser conocido. Este punto de partida es una abstracción teórica que debe de ser criticada para poder reconstruir la filosofía⁶.

Los problemas de la filosofía deben de ser re-examinados partiendo desde un punto de vista más radical y genuino. Los pragmatistas proponen que, en vez de partir desde una teoría o punto de vista teórico, debemos de comenzar en donde realmente estamos cada vez que filosofamos: dentro de nuestras vidas concretas. Esto es a lo que llaman James y Dewey la “experiencia”. Estar en la experiencia (o la vida) es estar íntimamente envuelto (como agente) en una interacción constante y no en una contemplación teórica y pasiva de objetos (como sujeto). Este es el sentido en que la “práctica” es primordial para el pragmatista.

⁴ Traducción de Peirce, CP5.416 por el Grupo de Estudios Peirceanos de la Universidad de Navarra. [En línea] <http://www.unav.es/gep/WhatPragmatismIs.html>.

⁵ Mi impresión de Peirce es el de un filósofo que aunque fue un genio, tenía una pierna en la tradición y otra en la manera nueva y radical de hacer filosofía.

⁶ Esto lo hace claro William James en su crítica al empirismo moderno. Pero la crítica de John Dewey es aún más comprensiva. Encontró que el error de empezar en el sitio inadecuado es tan común en la filosofía que decidió llamarlo, “la falacia filosófica.” Douglas Browning describe esta falacia y explica la crítica de Dewey al “el punto de partida teórico” en: BROWNING, Douglas. “Understanding Dewey: Starting at the Starting Point”. En: XIV Congreso Interamericano de Filosofía. Mexico, Agosto 19, 1999.

No todos los pragmatistas fueron suficientemente explícitos acerca de su punto de partida. A veces el entusiasmo con un descubrimiento bloquea la capacidad de articular de manera clara y precisa qué es lo que se ha encontrado. La mayoría de los pragmatistas simplemente adoptaron el nuevo punto de partida y procedieron a usarlo, haciendo filosofía que tratara de resolver y disolver problemas filosóficos. De todos ellos, Dewey es el más explícito. A través de sus escritos insistió en la importancia de tomar la experiencia como punto de partida. A pesar de esto, el tema del punto de partida ha sido ignorado por muchos de los pragmatistas contemporáneos. Esto en parte es culpa de Richard Rorty, quien ha sosteniendo que la preocupación de Dewey con la experiencia fue un grave error⁷. Más adelante indicaré por qué este rechazo de Rorty y otros está basado en un mal entendimiento sobre lo que es “la experiencia.”

Al igual que los pragmatistas clásicos, Ortega y Gasset y Risieri Frondizi examinaron cómo la filosofía moderna está basada en un punto de partida erróneo. El realismo y el idealismo moderno están en buen camino cuando buscan a una primera realidad pero desembocaron en sus posiciones problemáticas porque uno parte del objeto y otro del sujeto. El error es el mismo: “Ambas posiciones ofrecen un error común: haber exagerado uno u otro de los términos menospreciando la mutual relación. De ahí que frente a la “cosa en sí” del realismo haya surgido en los tiempos modernos “el sujeto en sí” del idealismo”⁸. En la filosofía moderna se empieza con la oposición entre el objeto y el sujeto. Asegura Frondizi que si en vez de este dualismo empezáramos con la experiencia en su integridad “desaparecen así los clásicos problemas sobre la posibilidad de que el sujeto pueda ponerse en relación con el objeto”⁹.

Los empiristas modernos (Locke, Hume) tenían buenas intenciones pero no supieron deshacerse del mismo punto de partida subjetivo de Descartes, donde el filósofo sólo tiene contacto con el contenido de su conciencia. El empirismo tradicional “se transformó automáticamente en racionalismo al volver la espalda a la realidad y encerrarse en sus propias construcciones”¹⁰. Las descripciones de los empiristas modernos del

⁷ Comparto con David Hildebrand su crítica a Rorty y a Hilary Putnam por no entender el punto de partida de Dewey, véase su *Beyond Realism and Antirealism*. Nashville, Tenn.: Vanderbilt University Press, 2003.

⁸ FRONDIZI. Op. cit., p. 49.

⁹ *Ibid.*, p. 102.

¹⁰ *Ibid.*, p. 122.

punto de partida no son más que fabricaciones teóricas que explican la experiencia pero no son la experiencia vivida. Por ejemplo, la descripción de una silla o una pared como un aglomerado de “*sense data*” o “simple ideas” no es más que una teoría sobre la silla y la pared, de seguro no son la silla y la pared que aparecen en mi vida (la experiencia vivida).

Si se describen las cosas tal y como son, veremos que la experiencia no es atomística (como pensaban los empiristas), que hay relaciones, y es todo un proceso. Si realmente y honestamente vamos a describir la experiencia tenemos que tratar de poner al lado las presuposiciones teóricas. Esta es exactamente la posición de William James y la razón por la cual caracterizó su empirismo como “radical”¹¹. Frondizi llama a su filosofía un “empirismo humanista”¹² pero es claro que es el mismo tipo de empirismo radical de James.

El empirismo adecuado es para Frondizi, uno que debe de “atenerse a las cosas mismas(...) pero no a las cosas desnaturalizadas por anticipaciones teóricas que tergiversen su sentido sino a ellas tal cual se presentan”¹³. Lo más importante es “evitar toda “idealización” o hipóstasis y partir de los fenómenos tal cual se nos presentan”¹⁴. Pero evitar esto ha sido muy difícil para los filósofos. La tentación del filósofo ha sido siempre la de tomar por hecho y como punto de partida lo que es sólo una construcción teórica. Dewey se dio cuenta de que aun James comete este error cuando describió la experiencia como algo “puro” en el sentido de algo anterior, aparte, o debajo de toda intervención o interpretación humana¹⁵. Pero la experiencia “pura” es sólo una explicación teórica de la experiencia y no una descripción fiel de la experiencia. En otras palabras, no es lo que encontramos de una manera inmediata y directa en nuestro diario vivir.

En esto Frondizi es claro y está de acuerdo con Dewey. La búsqueda del punto de partida es la búsqueda de algo dado, efectivo, inmediato y no puesto por nuestras teorías. Esta búsqueda ha sido resistida por algunos neo-pragmatistas. De hecho, han decidido abandonar la noción de la experiencia porque piensan que esto sería cometer el error filosófico conocido como “*the myth of the given*”. En una entrevista reciente Robert

¹¹ Cf. WILLIAM, James. *Essays in Radical Empiricism*. New York: Longman Green and Co, 1912.

¹² Cf. FRONDIZI, Risieri. “Bosquejo de mi filosofía: el empirismo como humanismo”. *En: Ensayos Filosóficos*. (s. f.)

¹³ FRONDIZI. *El Punto de Partida del Filosofar*. Op. cit., p. 8.

¹⁴ *Ibid.*, p. 40.

¹⁵ Dewey le hace esta crítica a James en: DEWEY, John. *Experience and Nature*. Op. cit., p. 40.

Brandom dice: *“Rorty and I both think that Sellars’ critique of the myth of the given shows the notion of experience as simply outmoded. I agree with him that there is no useful way to rehabilitate the concept of experience. We just need to do without that”*¹⁶.

Pero esto es un serio malentendido de la metafilosofía pragmatista. Tanto Dewey como Frondizi aclaran que exhortar a filósofos que deben de prestarle atención a la experiencia inmediata vivida no implica que sus filosofías sean epistemologías fundamentalistas basadas en la búsqueda de un dato “dado” en el sentido de un “‘hecho’ anterior o ajeno a toda interpretación”¹⁷. Negar un punto de partida absoluto o “puro” (libre de interpretación) no quiere decir que debamos abandonar la búsqueda de lo que es dado en nuestra experiencia inmediata del mundo. Todas nuestras experiencias presentes quizás están hasta cierto punto condicionadas por todo lo que heredamos (nuestra historia, cultura, carácter, y más), pero las experiencias vividas, y no las conclusiones de investigación, siguen siendo el “dato efectivo que hallemos en nuestra cuidadosa búsqueda y que *aprehendemos inmediatamente* por medio de un acto directo de nuestro espíritu”¹⁸.

La crítica de Frondizi va más allá de los filósofos modernos. Aunque admiraba a su contemporáneo Edmund Husserl criticó su punto de partida. El problema del método de Husserl (la “*epoche*”) es que la reflexión modifica la esencia de las vivencias y “reduce la vida a una pura contemplación de las esencias”¹⁹ pero la realidad es muy distinta. El método de Husserl pretende poner al lado nuestra participación en la experiencia como agentes. Husserl ignora lo que se podría llamar el aspecto “práctico” de la vida en su encuentro con el punto de partida²⁰. Dice Frondizi que en Husserl “poca importancia se otorga a la vida sentimental, valorativa y práctica –en oposición a la vida contemplativa– es una consecuencia de su actitud intelectualista”²¹. Este intelectualismo, que es tan común en la filosofía, es el resultado de asumir como punto de partida la vida descrita desde el punto de vista de un sujeto conocedor; que está separado y contemplando el contenido de su conciencia. Para Frondizi y los pragmatistas, esto equivale a partir

¹⁶ Interview with Robert Brandom. In: *Filosofisk Supplement*, (s.f.) p. 5

¹⁷ FRONDIZI. Op. cit., p. 49.

¹⁸ *Ibid.*, p. 51.

¹⁹ *Ibid.*, p. 77.

²⁰ Recientemente el pragmatista Douglas Browning ha hecho la misma crítica a Husserl en: BROWNING, Douglas. *Understanding Dewey: Starting at the Starting Point*. Op. cit., p. 8.

²¹ FRONDIZI. Op. cit., p. 77.

desde un punto de vista ajeno a la vida misma. Un empirismo “total” o “radical” requiere que el filósofo debe de partir a describir la realidad desde como ésta se nos revela cuando estamos haciendo, jugando, peleando, amando, gritando, y hablando.

III. DEFENSA DE LA EXPERIENCIA COMO EL PUNTO DE PARTIDA

El recuento anterior demuestra que los pragmatistas comparten una crítica fundamental a la filosofía moderna, pero ¿cuáles razones o argumentos han presentado para tomar la experiencia como el punto de partida adecuado de la filosofía? A manera de resumen, encontramos en los escritos de Dewey, Frondizi, James, y Ortega tres razones:

(a) La experiencia es la realidad más inmediata y neutral con respecto a nuestras teorías filosóficas.

(b) La experiencia es la realidad más radical en el sentido de ser inclusiva, primitiva, última, y forzada.

(c) La experiencia es la base constante, persistente, y confiable de evidencia en una investigación filosófica que es empírica.

(a) La experiencia es inmediata y neutral

Los pragmatistas rechazan el punto de partida moderno porque parten de lo que es una abstracción teórica, en vez de lo que es inmediatamente dado y presente en la experiencia. Si realmente se quiere “atenerse a las cosas mismas tal cual se nos presentan” tenemos que admitir que ningún filósofo empieza a filosofar metido dentro de su conciencia. Un empirismo total y honesto reconoce que empezamos siempre en un contexto particular donde estamos (el yo) entre las cosas, envuelto en actividades dirigidas hacia ciertos objetos, somos organismos en un medio ambiente y en interacción con otras cosas. Esto es lo que el pragmatista llama la “experiencia”. A diferencia de la concepción moderna, la experiencia no es algo subjetivo y cognoscitivo, ni tampoco es el confrontamiento de un yo, una actividad, y un objeto como entidades que de alguna manera tienen una existencia anterior. Se empieza con la relación (interacción) entre “el yo” y los objetos en el mundo y no con un dualismo.

Sólo la experiencia es algo “dado”, “efectivo”, e “inmediato” y no puesto por nuestras teorías. No es la conclusión de investigación. El que este “dato” no provenga de una teoría en particular significa que antes de ser examinado tiene un tipo de neutralidad con respecto a las diferentes teorías que se puedan articular sobre la experiencia. Para ser empírico en la filosofía hay que tener mucho cuidado. Debemos evitar la confusión entre lo que encontramos y lo que queremos encontrar debido a nuestros prejuicios teóricos. Por ejemplo, ¿qué hace “moral” a un problema? Si vamos a investigar esto deberíamos tratar de poner nuestros prejuicios teóricos a un lado y tratar de designar, ubicar, apuntar y luego describir tantos problemas que sufrimos en nuestras vidas se nos presentan como morales y no estéticos o económicos. Esto no es fácil para el filósofo. Por ejemplo, un filósofo utilitarista que este más interesado en la validez de su teoría que en honestamente investigar la realidad moral y vivida (tal y como ésta se nos aparece), tendrá la tentación de describir todos los problemas morales en términos de un conflicto de unidades de utilidad. Si es así, confunde su descripción teórica con la experiencia inmediata.

(b) La experiencia es la realidad radical

La filosofía moderna asume que la investigación filosófica es algo que ocurre y empieza dentro de un plano subjetivo (el contenido mental) que intenta conocer una realidad independiente y no cambiante. La alternativa a este punto de partida “subjetivo”, por lo general, ha sido uno “objetivo” que consiste en adoptar un “punto de vista de Arquímedes” o “punto de vista de Dios”. Muchos filósofos “post-modernos” argumentan que estos dos puntos de vista (o puntos de partida) son ilusorios porque la investigación filosófica está siempre en una historia, en una cultura, en una lengua, o en una sociedad particular. Pero de acuerdo a los pragmatistas ninguno de estos contextos constituyen el contexto que en última instancia estamos cuando filosofamos. Como Pierce insistió, empecemos “donde estamos” y siempre estamos, en una situación. Todas las realidades discutidas por los filósofos (incluyendo el lenguaje, esquemas conceptuales, y la cultura) se encuentran en situaciones presentes, concretas y vividas, en otras palabras, en la experiencia. Dice Frondizi, “La experiencia constituye la realidad o hecho fundamental puesto que todos los demás “hechos”, y todas las demás formas de “realidad”, se dan en ella como sus miembros, partes o elementos”²². Nada hay antes o por encima de la experiencia. Esto fue

²² *Ibid.*, p. 85.

lo que condujo a José Ortega y Gasset a decir, que “mi vida” (término equivalente a la “experiencia”) es una realidad “radical” en el sentido de que “es la raíz de todas las demás en el sentido de que estas, sean las que fueren, tienen, para sernos realidad, que hacerse de algún modo presentes o, al menos, anunciarse en los ámbitos estremecidos de nuestra propia vida”²³.

Esto no niega que a lo mejor hay realidades fuera de la experiencia, pero no hay manera de llegar a ellas excepto a través de nuestras experiencias vividas en situaciones, desde nuestra piel. Quieran o no lo quieran, todos los filósofos han empezado en la experiencia. Por esto es que Frondizi afirma que la experiencia es no sólo “la fuente legítima” sino que “el punto obligado de referencia de la teoría filosófica”²⁴. Por lo tanto, la única diferencia que existe entre filósofos es entre aquellos que lo niegan y aquellos que lo admiten y tratan de hacer filosofía basada en este hecho fundamental. Descartes afirmó que su punto de partida fue en su mente (lo que apareció en su conciencia), pero sabemos cuál fue realmente su punto de partida. Mientras meditaba estaba al frente de un fuego, entre muebles y sin abandonar todos los prejuicios que trataba de negar.

Un punto importante que Ortega destaca acerca de Descartes es que tanto el fuego como las paredes de su casa, estaban presente como realidad inmediata y no, como suponía Descartes, eran apariencias u objetos de su conocer o de creencia. La realidad se tiene, se presenta, que es diferente a decir que se conoce. Este es el truco moderno que todavía existe en la filosofía, se piensa que para ser realista hay que justificar como conocimiento verdadero creencias sobre la realidad. El conocer es sólo una de tantas relaciones que podemos tener con la realidad. Esa pared es real, está presente, aun si no es objeto de mi conocimiento. “Esa pared está ahí” es fundamental y no es lo mismo que decir “Yo conozco o creo que esa pared está ahí”.

(c) La experiencia como base de evidencia de la investigación filosófica

A pesar de lo equivocado que estaba Descartes, entendió la importancia del problema del punto de partida para la filosofía. Pensó que si encontramos y hacemos claro cual debe de ser el punto de partida entonces podemos hacer filosofía guiada por ese punto de partida.

²³ ORTEGA Y GASSET, José. *El Hombre y La Gente*. Madrid: Revistas de Occidente, 1957. p. 47.

²⁴ FRONDIZI. *Op. cit.*, p. 12.

En otras palabras, el asunto no es sólo dónde empezamos sino cómo procedemos al hacer filosofía. El punto de partida es también dónde regresamos para apoyar nuestras teorías. La realidad que sirve de sostén a las construcciones teóricas también sirve como la base de donde buscar apoyo y evidencia. Como la experiencia es una realidad cambiante, la filosofía tiene que tomar en serio un regreso constante a la experiencia.

El retorno constante de nuestras teorías e investigaciones a la vida diaria es la única manera de tratar de prevenir los vicios de la filosofía, especialmente la esterilidad de que a veces se nos acusa. No nos debe de sorprender que cuando empezamos con abstracciones que no corresponden a la experiencia de nadie, terminamos con filosofías que tienen poca relevancia en nuestras vidas. Es importante para Frondizi “volver-una vez-más-al punto de partida. Esforzarse por alcanzar la realidad que sirve de sostén a las construcciones” de otra manera la filosofía pierde vigencia. Esta es la única manera de impedir “que la filosofía se convierta en un juego intelectual estéril y ajeno a las preocupaciones irrenunciables de la vida humana.”

Para los pragmatistas la experiencia es el punto de partida adecuado de toda investigación filosófica. Además, los pragmatistas sostuvieron que es posible hacer una descripción de la experiencia que sea fiel a la experiencia vivida y esté libre de toda teoría. A continuación me propongo concluir este ensayo presentando algunas de sus descripciones generales.

IV. DESCRIPCIÓN GENERAL DE LA EXPERIENCIA

(a) La integridad de la experiencia

Según James y Dewey la filosofía moderna empezó con divisiones ontológicas (dualismos) en vez de empezar con la “unidad integral” que encontramos en la experiencia. Es por esto que terminan creando problemas de cómo unir y relacionar lo que ellos han separado por medio de una abstracción. Este es el error de la concepción subjetiva de la experiencia. Podemos distinguir entre el que percibe y lo percibido pero ni el uno ni el otro son antecedentes a la experiencia, son partes de una unidad integral. Como bien explica Dewey la experiencia tiene dos caras sin perder su integridad. *“It is “double-barrelled” in that it recognizes in its primary integrity no division between act and material, subject and*

object, but contains them both in an unanalyzed totality"²⁵. Frondizi dice algo muy similar: "La experiencia ofrece la aparente paradoja de que está condenada a ser mía – se da siempre en primera persona – y al mismo tiempo que necesita de lo que no es mío para existir. Tiene una cara subjetiva y otra objetiva"²⁶.

Dewey y Frondizi no siempre comparten la misma terminología. Dewey describe la experiencia en términos de un organismo en un medio ambiente pero hace claro que la distinción no implica separación ontológica. Frondizi describe la experiencia en términos del yo, mis actividades y los objetos, pero hace el mismo tipo de aclaración acerca de la integridad de la experiencia. "Los miembros que constituyen la experiencia" se nos dan "como una totalidad indivisible"²⁷. En otras palabras, "los tres miembros *no preexisten* a su relación sino que *se constituyen* en ella"²⁸. Son "miembros de una totalidad única en la que las relaciones son tan importantes como los elementos. El yo es lo que es en tanto ejercita una actividad sobre un objeto. Este, a su vez, no existe como objeto con anterioridad a su relación con aquel sino que es objeto de la actividad del yo"²⁹.

(b) El aspecto social de la experiencia

Todos los pragmatistas tomaron como inaceptable la noción de que el empezar con la experiencia es el empezar dentro de mí, como si otros y lo social fuera algo secundario. La existencia de otros es un "pseudoproblema" que se crea cuando se cree empezar a filosofar encerrados en nuestra conciencia.

Frondizi afirma lo mismo y critica la teoría de "razonamiento por analogía"³⁰. El conocimiento de otros es igual de inmediato y primario, no es la conclusión de una inferencia. "El yo ajeno no se nos da, por cierto, como un objeto entre los demás objetos. No percibimos un cuerpo y luego inferimos la existencia de una psiquis que lo anima, sino que lo aprehendemos desde un principio como una unidad semejante a la de nuestro yo"³¹. En el punto de partida "compartimos el mundo

²⁵ DEWEY. *Experience and Nature*. Op. cit., p. 18.

²⁶ FRONDIZI. Op. cit., p. 116.

²⁷ *Ibid.*, p. 84.

²⁸ *Ibid.*, p. 128.

²⁹ *Ibid.*, p. 125.

³⁰ *Ibid.*, p. 108-109.

³¹ *Ibid.*, p. 110.

objetivo,” nos encontramos con las actividades de otros seres, y algunos son semejantes a nosotros. Dice Frondizi que “otros son partes de mi experiencia”; “nuestra experiencia es inconcebible sin la experiencia ajena”; “vivir equivale a convivir”; “vivimos en comunidad, si perder por eso nuestra individualidad personal”³². Como si fuera poco, Frondizi afirma que “el conocimiento del yo ajeno es simultáneo al autoconocimiento y al mundo exterior”³³. Aunque Frondizi no elabora esta idea, es claro que considera razonable una teoría del conocimiento del “Yo” muy similar a la del pragmatista George Herbert Mead³⁴.

(c) La experiencia como proceso en el presente y con continuidades

Todos los pragmatistas dicen que la experiencia es un proceso fundamentalmente anclado en una situación presente que es continua con previas situaciones presentes, y que estamos siempre en el presente. “Hablar de una vivencia presente en un instante es una abstracción y equivale a detener lo moviente, a matar la vida”³⁵. Tanto el “yo” como toda la realidad es un proceso y no una substancia. Los pragmatistas encuentran el aspecto fluyente de la realidad compatible con afirmar también que hay continuidad y estabilidad. “La naturaleza fluyente y temporal de la experiencia no implica la negación de su unidad”³⁶. “El carácter fluyente de la experiencia no es incompatible con la presencia de elementos estables”³⁷.

La experiencia es un proceso anclado siempre en el presente, donde lo que llamamos pasado y futuro son aspectos de un presente en movimiento. El pasado es para los pragmatistas memorias presentes y el futuro son anticipaciones que ocurren en el presente. George Herbert Mead defiende esta noción del tiempo en *The Philosophy of the Present*³⁸. De la misma manera Frondizi afirma que, el “ahora y aquí” como un pedazo inmóvil es una abstracción “del carácter temporal y fluyente de la experiencia” presente que incluye “el sedimento de las experiencias pasadas y la proyección hacia el futuro”³⁹. Por lo tanto,

³² *Ibid.*, p. 116.

³³ *Ibid.*, p. 108.

³⁴ Mead presenta una teoría del conocimiento del “yo” como resultado de interacción social en: MEAD, George Herbert. *Mind, Self, & Society*. Chicago: The University of Chicago Press, 1934.

³⁵ FRONDIZI. Op. cit., p. 96.

³⁶ *Ibid.*, p. 117.

³⁷ *Ibid.*, p. 124.

³⁸ MEAD, George Herbert. *The Philosophy of the Present*. Murphy, Arthur E. (Ed.). Chicago: University of Chicago Press, 1932.

³⁹ FRONDIZI. Op. cit., p. 114.

“Mi experiencia aquí y ahora es el hecho primario e innegable, el punto inicial que buscábamos”⁴⁰.

Este punto de partida es muy diferente a la concepción del tiempo donde el presente, el pasado, y el futuro son espacios o sitios en una línea vista desde un punto de vista removida del tiempo. Esta es una concepción del tiempo desde un punto de vista contemplativo y teórico. Esta noción del tiempo nos ha sido muy útil para organizar nuestras vidas y para las ciencias pero no es el tiempo vivido desde adentro de la vida como agente.

(d) La diversidad y el aspecto práctico de la experiencia

La noción tradicional de la experiencia reduce toda la experiencia a las actividades relacionadas con el conocimiento. Esto es consecuencia de tomar como el punto de partida de la filosofía lo que se revela a la conciencia de un sujeto como conocedor. Pero un empirismo “total” reconoce la riqueza o variedad en que se da la experiencia.

Nuestra participación primordial en la experiencia es como agente y no como sujeto conocedor. El “yo haciendo” es mucho más fundamental que el “yo conociendo” ya que el conocer es sólo una de las cosas que “hacemos”. Pero debido al intelectualismo en la filosofía que reduce las actividades del yo a las de orden intelectual exclusivamente, la realidad se ha convertido en sólo un “objeto de conocimiento”. Este es un prejuicio que tiene poca base en la experiencia vivida.

e) La experiencia es situacional

Ya hemos mencionado este aspecto clave de los pragmatistas. Dewey, especialmente, enfatiza que el estar en la experiencia es el estar siempre en una situación. Una situación es el contexto particular, único, y cualitativo donde se encuentran el organismo y su medio ambiente. Dewey describe la experiencia en términos de lo que hace y lo que le sucede a un organismo en una situación. El conocimiento es una actividad o proceso que ocurre en una situación. Es en efecto, el resultado de una investigación que ha transformado la calidad de una situación indeterminada o dudosa.

⁴⁰ *Ibid.*, p. 124.

Aunque he sido breve en presentar algunas de las descripciones generales de la “experiencia”, creo que es importante para evitar mal entendimientos comunes sobre el pragmatismo. Queda claro que su metafilosofía es muy diferente a la asumida (como no problemática) por los filósofos empíricos modernos. Pero es importante también reconocer cuan diferente es su metafilosofía a la asumida por algunos filósofos hoy en día que se consideran post-modernistas y neo-pragmatistas. Creo que es importante establecer un diálogo genuino donde se discutan estas diferencias.

REFERENCIAS

- BROWNING, Douglas. (Agosto 19 de 1999). “Understanding Dewey: Starting at the Starting Point”. En: XIV Congreso Interamericano de Filosofía. Mexico.
- DEWEY. John. (1925). *Experience and Nature Later Works*. Southern Illinois University Press, Vol. 1.
- FRONDIZI, Risieri. (s.f.). “Bosquejo de mi filosofía: el empirismo como humanismo”. En: *Ensayos Filosóficos*.
- _____. (1945). *El Punto de Partida del Filosofar*. Buenos Aires: Editorial Losada. 2a. Ed., 1957.
- HILDEBRAND, David. (2003). *Beyond Realism and Antirealism*. Nashville, Tenn.: Vanderbilt University Press.
- MEAD, George Herbert. (1934). *Mind, Self, & Society*. Chicago: The University of Chicago Press.
- _____. (1932). *The Philosophy of the Present*. Murphy, Arthur E. (Ed.). Chicago: University of Chicago Press.
- ORTEGA Y GASSETT, José. (1957). *El Hombre y La Gente*. Madrid: Revista de Occidente.
- WILLIAM, James. (1912). *Essays in Radical Empiricism*. New York: Longman Green and Co.

LA NEUTRALIDAD VALORATIVA DE LA CIENCIA Y EL CONFLICTO ENTRE LA RAZÓN Y LAS PASIONES

SCIENCE'S NEUTRALITY AND THE CONFLICT BETWEEN REASON AND PASSIONS

CARLOS EMILIO GARCÍA

Universidad de Caldas, Colombia. carlos.garcia_d@ucaldas.edu.co

PABLO R. ARANGO

Universidad de Caldas, Colombia. pablo.arango_g@ucaldas.edu.co

RECIBIDO EL 27 DE AGOSTO DE 2010 Y APROBADO EL 30 DE SEPTIEMBRE DE 2010

RESUMEN ABSTRACT

El presente trabajo discute, en primer lugar, la tesis de la neutralidad moral de la ciencia. Se examinan, a la luz de los puntos de vista del Positivismo Lógico, los posibles significados de esta tesis, así como la plausibilidad de cada uno. En segundo lugar, el trabajo analiza el problema de las relaciones entre la ciencia y el mal. Apelando a algunas tesis de David Hume, se argumenta que el problema no reside en la ciencia como tal, sino más bien en la parte irracional de la naturaleza humana. Se concluye que lo que se necesita para enfrentar las enormes amenazas destructivas que surgen de ciertas aplicaciones de la ciencia, no es un cambio en el aspecto racional del hombre, sino justamente un cambio en su parte irracional.

This essays discusses, firstly, the thesis of science's neutrality. In the light of the views of logical positivism, the possible meanings of this thesis are examined, as well as their plausibility. Secondly, we analyze the problem of the relationships between evil and good. Appealing to certain Humean views, we argue that the problem does not lie in science itself, but in the irrational side of human nature. We conclude that what is needed to face the big destructive threats that come from certain applications of science is not a change in the rational side of man, but precisely a change in its irrational side.

PALABRAS CLAVE KEY WORDS

bien, ciencia, ética, Hume, mal, neutralidad valorativa, pasiones, razón.

good, science, ethics, Hume, evil, science neutrality, passion, reason.

*“La ciencia es admirable porque nos otorga poder sobre la naturaleza,
y dicho poder proviene completamente de la técnica”.*

Bertrand Russell

INTRODUCCIÓN

La discusión sobre la responsabilidad moral de la ciencia (o de los científicos) reaparece con cierta frecuencia en la literatura especializada sobre ética y epistemología. Las respuestas varían entre el rechazo rotundo de que tal tipo de responsabilidad sea atribuible en el contexto de las consecuencias pragmáticas, la aceptación con matizaciones, y la afirmación *simpliciter* que atribuye responsabilidad moral a los científicos y se apresura a extenderla, sin ninguna cualificación adicional, a esa abstracción que denominamos “ciencia”. En este ensayo, nos proponemos examinar la tesis de la neutralidad valorativa de la ciencia, distinguir entre algunas de sus formulaciones (con énfasis en las que plantearon los positivistas lógicos), revisar brevemente las críticas a esta posición, y estudiar su aplicación a la evaluación de un hecho histórico que se ha convertido en ejemplo recurrente: la fabricación y posterior detonación de la bomba atómica (cuyo sexagésimo quinto aniversario se conmemoró recientemente). Para este análisis, hemos considerado pertinente discutir una confrontación, quizá tan notoria como la de la neutralidad valorativa de la ciencia; a saber, la de la supuesta tensión entre la razón y las pasiones. Como bien se sabe, suele argüirse que si los seres humanos fuéramos perfectamente racionales, episodios como el holocausto nazi, o las guerras mundiales, jamás habrían ocurrido. Acudiendo a las tesis de Hume, argumentaremos que el problema fundamental no se resuelve en el nivel de la racionalidad, sino más bien en el de las emociones. En las secciones siguientes, abordaremos estos temas en el mismo orden en que los hemos presentado.

1. LA NEUTRALIDAD VALORATIVA DE LA CIENCIA

Cualquier persona medianamente culta comprende la expresión “la ciencia es éticamente neutral”. Por lo general, este planteamiento tiende a interpretarse como un llamado a diferenciar entre las realizaciones prácticas de la ciencia y las consecuencias que pueden traer su aplicación para la naturaleza o el hombre. En términos más coloquiales, se supone que la idea desvela un claro contraste, algo del tipo: “una cosa son los productos de la ciencia, y otra muy distinta las consecuencias sociales

o materiales de su uso indebido, precipitado o indiscriminado"; o, en algunos casos, una afirmación más precisa como: "en razón del carácter público de la ciencia, una vez se difunde un descubrimiento particular, el científico no tiene responsabilidad por el uso que otros le den a sus inventos". Al parecer, la intención subyacente en las proposiciones anteriores consiste en señalar el hecho incontrovertible de que los creadores son incapaces de prever (y mucho menos de controlar) todas las consecuencias de sus creaciones. Del mismo modo que no es posible predecir las múltiples formas de usar indebidamente cualquier objeto (natural o artificial), no resulta dable vaticinar el sino de una realización técnica. Expresada de esta manera, la tesis es verdadera en sentido trivial. *Mutatis mutandis*, no sólo los científicos, sino todos los demás creadores son incapaces de anticipar todos los usos posibles (incluyendo aquellos que puedan resultar nefastos o moralmente cuestionables) de los productos de su ingenio.

La tesis de la neutralidad valorativa de la ciencia se atribuye a los positivistas lógicos. Sin embargo, hay cierta ligereza en esta atribución. Si examinamos los artículos que publicaron en torno al tema de la ética sus principales representantes, todo lo que encontramos es una discusión sobre el estatus cognoscitivo de esta disciplina, una indagación acerca de la naturaleza de los enunciados prácticos y un examen sobre la cuestión de si es posible tratar sus proposiciones con las mismas herramientas con que ellos deseaban analizar los enunciados de la ciencia. Así, por ejemplo, Schlick incluyó la ética en las ciencias sociales. Convencido de que los resultados de la ética no dependían del uso de ninguna facultad especial de intuición moral, sostuvo que las cuestiones éticas eran cuestiones de hecho del tipo: por qué la gente sostiene ciertos principios, qué es lo que desean, cómo pueden satisfacer sus deseos, etc., y concluyó que la ética, al igual que la ciencia, busca el conocimiento y encuentra la fuente material de sus ideas en la experiencia¹.

Algo similar puede inferirse del análisis lingüístico que hace Stevenson. Según su opinión, la dificultad para resolver las cuestiones morales se deriva de la ambigüedad y confusión que usualmente acompaña los

¹ Desde este punto de vista, la ética no es otra cosa que teoría del conocimiento, cuya tarea no puede ser producir o establecer la moralidad. "[La ética] no tiene la tarea de producir el bien -ni en el sentido de que le competa investir de realidad el bien en los asuntos humanos, ni en el sentido de que deba estipular o decretar lo que debe significar la palabra 'bien'. No crea el concepto ni los objetos que caen bajo el concepto, ni proporciona oportunidades para aplicar el concepto a los objetos". SCHLICK, Moritz. "What is the Aim of Ethics?" In: AYER, Alfred (Ed). *Logical Positivism*. New York: The Free Press, 1959. p. 248.

términos éticos. Así por ejemplo, la palabra “bueno” se ha definido con ayuda de la noción de aprobación (aquello que la mayoría aprueba) o apelando a ciertas actitudes psicológicas (lo que yo deseo). Pero semejante dualidad impide el uso unívoco del término, cuya acepción más fuerte debería satisfacer nuestras intuiciones filosóficas al menos en tres aspectos: (a) la posibilidad de tener (y decidir) desacuerdos genuinos sobre la bondad de algo; (b) la fuerza pragmática que nos llevaría a proceder favoreciendo aquellos actos que hemos calificado como buenos; y (c) la posibilidad de verificar la “bondad” de algo, acudiendo a métodos extra-científicos. Stevenson concluye que para tratar de manera adecuada este problema, es necesario renunciar a las teorías éticas tradicionales y sustituirlas por una variante menos dependiente de la descripción de nuestros intereses, y capaz de explicar la forma como los enunciados éticos son capaces de influir sobre la conducta de nuestros semejantes. De este modo, cuando se le dice a alguien que la sociedad condena el robo, y que, en consecuencia, debe abstenerse de robar, el juicio moral “tiene una fuerza cuasi-imperativa que, al operar mediante la sugestión, e intensificada por su tono de voz, fácilmente le permite a usted comenzar a *influir* o *modificar* los intereses [del interlocutor]”². Los enunciados de la ética son, pues, medios para presentar nuestras sugerencias de modo más eficaz, y si bien es posible determinar empíricamente qué tan influyentes resultan, una parte importante de los usos de lenguaje moral depende de las intuiciones que hemos formado en virtud de la influencia de las ideas prevalecientes en las sociedades en que hemos sido criados.

No consideramos necesario abundar en ejemplos para mostrar que la tesis de la neutralidad valorativa de la ciencia no es muy frecuente en los textos de los autores positivistas. Sin embargo, es preciso reconocer que dicha tesis puede derivarse de los planteamientos centrales de la doctrina y, sobre todo, de las conclusiones que los positivistas propusieron sobre la naturaleza de la ciencia y de la ética, y de los resultados de sus análisis sobre las particularidades del lenguaje moral. Un caso ilustrativo se puede encontrar en un ensayo de Weber sobre las implicaciones de esta misma tesis para las ciencias sociales. El autor formula la tesis que estamos discutiendo en un lenguaje claramente positivista, mediante el cual invita al lector a considerar la aceptabilidad y el significado de un principio lógico —según el cual no hay una ruta que conecte directamente los estados de cosas y los valores— subyacentes a la

² STEVENSON, C. L. “The emotive Meaning of Ethical Terms”. In: Logical Positivism. *Ibid.*, p. 269.

pregunta de “si una ciencia particular está ‘libre’ de juicios de valor (...)”³. Es evidente, según este fragmento, que Weber tiene en mente el problema de las relaciones entre ética y ciencia, al igual que la cuestión de decidir si los investigadores deben mantener separadas las esferas del mundo de los hechos y de los valores, y si tal proceder confirmaría la supuesta neutralidad valorativa de la ciencia. En la sección siguiente, expondremos los argumentos más importantes de la discusión entre positivistas y anti-positivistas.

2. LA CONTROVERSIA SOBRE LA NEUTRALIDAD VALORATIVA DE LA CIENCIA

Durante el período en que el positivismo lógico estuvo en el centro del interés intelectual, solía discutirse con fruición la tesis de la neutralidad valorativa de la ciencia. Los críticos del positivismo, deseosos de extender sus objeciones al conjunto de las implicaciones de la doctrina, se apresuraban a señalar el hecho —igualmente incontrovertible— del carácter nefasto de actos que sólo resultaron posibles debido a realizaciones científicas muy particulares. La metáfora de la caja de Pandora les servía para matizar sus puntos de vista. Piénsese, por ejemplo, en inventos como la dinamita, la bomba atómica, o el gas sarín (también llamado mostaza), para no citar sino tres. Resulta evidente, no sólo que inventos de esta naturaleza han sido usados de manera moralmente reprobable sino que, una vez desarrollados, pasaron a constituirse en una amenaza real para la supervivencia de la raza humana y para la estabilidad de la civilización (incluyendo en ella a la ciencia) que los hizo posible. Por otra parte, los hechos de la historia han terminado por justificar la suspicacia con que los críticos ven (y censuran) ciertas aplicaciones prácticas de inventos como los que acabamos de mencionar. Nadie puede negar que estas creaciones han servido a propósitos moralmente cuestionables en manos de gobiernos, grupos o individuos venales o aviesos en todo el planeta. Por razones como estas, los enemigos del positivismo lógico creyeron tener un caso sólido para demostrar la falsedad de la tesis de la neutralidad valorativa de la ciencia. ¿Cómo puede ser éticamente neutral una ciencia que amplifica el poder

³ WEBER, Max. “Value Judgments in Social Science”. In: BOY, Richard *et. al* (Eds). *The Philosophy of Science*. Cambridge: The MIT Press, 1991. p. 719. El autor insiste en que la expresión “juicio de valor” ha de entenderse desde la perspectiva de las evaluaciones prácticas de aquellos fenómenos que pueden ser influidos por nuestras acciones, y calificados como dignos de aprobación o condena.

del ser humano para destruir y hacer daño, y algunos de cuyos logros prácticos han infligido dosis considerables de dolor y sufrimiento?⁴

Por su parte, los positivistas lógicos acudieron a un expediente bien distinto para responder a estas críticas y defender la tesis de la neutralidad valorativa de la ciencia. En primer lugar, fieles al ideal romántico sobre el carácter emancipador del conocimiento científico (inaugurado desde los albores de la modernidad), se esforzaron por mostrar que la ciencia es, ante todo, la mejor forma de resolver los problemas más urgentes de la sociedad, de mejorar significativamente la calidad de la vida de los asociados y, como consecuencia de esto, de conseguir una sociedad más feliz, justa e ilustrada. Como pruebas de semejante potencialidad, basta con considerar la influencia de logros como la agricultura, o la medicina moderna sobre la calidad de vida de las sociedades avanzadas. Lo malo –repetían– no son los inventos obtenidos gracias al grado creciente de nuestro dominio sobre la naturaleza, sino el uso que algunos individuos –ya sea que obren de manera aislada o en grupo, a nombre propio o escudados bajo la protección de una sociedad o una ideología particular– hacen de tales realizaciones⁵. Desde este punto de vista, culpar a Einstein o a Openheimen de las consecuencias devastadoras de las bombas lanzadas sobre Hiroshima y Nagasaki, resulta tan inadecuado como culpar a Flaubert, por los suicidios de muchas mujeres infelices que quisieron seguir el ejemplo de Madame Bovary para poner término a su desolación. Dicho de otro modo, si bien hay un punto de contacto en alguno de los eslabones de la extensa cadena causal que lleva de las teorías físicas hasta la fabricación y posterior detonación de la bomba atómica, este punto es tan tangencial y lejano, que sólo mediante un razonamiento sumamente defectuoso podría resultar factible culpar directamente a los científicos citados (o a otros) de tan desafortunados hechos.

⁴ Es importante tener en cuenta que, si nos concentramos en el aspecto teórico de la ciencia, sería necesario modificar sensiblemente la acusación, pues otros sistemas de ideas son igualmente culpables de crímenes similares. Piénsese, por ejemplo, en todas las consecuencias desastrosas de la religión. Los actos inspirados por el fanatismo religioso han ocasionado casi tantos males y sufrimiento como cualquiera de los inventos ‘diabólicos’ de la ciencia.

⁵ Por otro lado, incluso autores como Popper reconocen el contraste entre las consecuencias deseables y nefastas de la ciencia, cuyo progreso “(...) – que constituye en parte una consecuencia del ideal de autoemancipación por el conocimiento – está contribuyendo a la prolongación y enriquecimiento de nuestras vidas; pero nos ha llevado a atravesar la vida con la amenaza de una guerra atómica, y es dudoso que, a fin de cuentas, haya contribuido a la felicidad y el contento del ser humano”. POPPER, Karl. En busca de un mundo mejor. Barcelona: Paidós, 1996. p. 184.

Discusiones más recientes sobre este tópico, introducen un elemento que fue insuficientemente analizado en el debate que estamos reseñando; a saber, la naturaleza de las relaciones entre ciencia y política, junto con el problema de la responsabilidad social del científico. Actualmente, se acepta que la sociedad debe ejercer alguna forma de control sobre la ciencia y los científicos⁶. Al fin y al cabo, al menos en los países desarrollados, la mayor parte de las actividades científicas se llevan a cabo con la financiación de dineros públicos, por lo que resulta apropiado que el Estado, en representación de los contribuyentes, supervise y controle tales actividades. Por otra parte, es difícil negar que los científicos (como otros trabajadores de la cultura) tienen responsabilidades sociales ineludibles y que, en consecuencia, no pueden estar exentos de ciertas formas mínimas de supervisión y control. De hecho, pensadores como Feyerabend han llevado al extremo la idea de que la sociedad puede y debe controlar a los científicos (y como resultado obvio, a la ciencia misma). Según sus planteamientos, no es ni siquiera necesario que los encargados de ejercer tales formas de control sean ellos mismos expertos en ciencia. A partir de la tesis de que en una sociedad democrática debe haber una separación entre ciencia y sociedad, paralela a la separación que hay entre religión y ciencia (al menos en los países no confesionales), Feyerabend sostiene que los legos pueden ser representantes genuinos de los intereses de la mayoría, y que en tal condición están en capacidad de decidir (e imponer sus decisiones sobre los científicos) lo que resulte más conveniente para los intereses de la sociedad⁷.

No vamos a discutir detalladamente la teoría feyerabendiana sobre el rol de la ciencia en una sociedad genuinamente democrática, pues este tema no forma parte esencial de nuestros propósitos en este ensayo. Con todo, vale la pena tomar en cuenta que este pensador vienés logró identificar correctamente al menos una de las estrategias mediante las cuales se pueden controlar o minimizar los efectos nocivos que, para la salud de la sociedad, tienen algunas aplicaciones concretas de la ciencia. Independientemente de cuál sea la respuesta al problema de la neutralidad ética de la ciencia, resulta alentador descubrir que el control político y la presión de los ciudadanos pueden detener, o cuando menos morigerar, las consecuencias negativas que el uso de ciertos inventos puede desencadenar. Por otra parte, resulta indispensable

⁶ Decisiones como la de prohibir cualquier tipo de experimento relacionado con la clonación de seres humanos, o la de limitar la investigación sobre "células madre" provenientes de embriones humanos, corresponden a formas claras de control político sobre la actividad de los científicos.

⁷ Ver, por ejemplo, su *Ciencia en una Sociedad Libre*. Madrid: Siglo XXI, 1987.

señalar algunos de los defectos de la discusión sobre este problema. Para repetirlo una vez más, la tesis es verdadera en sentido trivial, cuando se la interpreta como si se refiriera a una abstracción (la ciencia misma), pero nuestras intuiciones morales no nos dejan dudas sobre la responsabilidad moral que compete a cualquier científico que participa consciente y voluntariamente en procesos de investigación destinados a desarrollar artefactos o productos concebidos deliberadamente con el fin de causar muerte y destrucción. En lo que sigue, intentaremos determinar la posición correcta sobre la tesis de la neutralidad valorativa, y considerar, desde una perspectiva más pragmática, las relaciones entre ciencia y maldad.

3. CIENCIA, TECNOLOGÍA Y MAL

Las relaciones entre la ciencia, la tecnología y el mal son complicadas. De no haber sido por los desarrollos en física subatómica, las bombas nucleares jamás habrían existido. Como ocurre con todos los contrafácticos, éste expresa una verdad vacía; pues de modo semejante, tampoco habrían sido posibles ciertos avances en medicina nuclear que actualmente apreciamos enormemente. Sin embargo, las bombas nos ponen hoy frente a la posibilidad, siempre al alcance de la mano, de una eliminación casi total de la vida humana. Éste es sólo un caso, pero uno de los más perspicuos. ¿Cuál es, entonces, la relación entre ciencia, tecnología y maldad? Algunas ideas de lo que podemos denominar, vagamente, el clima intelectual de nuestra época, sugieren que la raíz de muchos de los problemas morales que plantea el desarrollo científico y tecnológico debe ser buscada en la concepción racionalista de la modernidad. Según esta visión, hay algo profundamente equivocado en el ideal de educar a los hombres para que sean, ante todo, seres racionales, y valdría la pena sustituir este ideal por tradiciones intelectuales menos rigurosas.

Nuestro punto de vista se opone radicalmente a estas ideas. Para explicarlo, recurriremos a las tesis de uno de los filósofos modernos más importantes: David Hume. Famosamente, Hume escribió que “la razón es, y debe ser, esclava de las pasiones, y no puede pretender otro oficio que el de servir las y obedecerlas”⁸. Haciéndonos eco de Hume, sostenemos que, por lo que respecta a las fuentes de donde emana la

⁸ HUME, David. *Tratado de la Naturaleza Humana*. Madrid: Editora Nacional, 1981. p. 617.

maldad, el problema no está en la “razón pura”, sino más bien en las pasiones. La tesis de Hume es que el motor de las acciones nunca es la parte puramente intelectual del ser humano, sino precisamente su parte emocional. Si uno no tiene ningún deseo específico, la razón no puede decirle nada acerca de lo que debe hacer o conviene evitar. En otras palabras, la deliberación que tiene lugar en el ámbito de lo que los teóricos denominan “razón prudencial” sólo puede ser iniciada si median deseos definidos y el sujeto necesita decidir si los sigue o los acalla. Bertrand Russell ilustra muy bien este punto:

Los enemigos de la razón (...) piensan que los defensores de la racionalidad quieren que la razón dicte los fines al igual que los medios (...) Cuando se me dice, como ocurre con frecuencia, que apenas tengo en cuenta el papel que juegan las emociones en los asuntos humanos, me pregunto qué fuerza motora supone el crítico que considero dominante. Los deseos, las emociones, las pasiones (...) son las únicas causas de la acción. La razón no es causa de la acción sino sólo un regulador. Si yo deseo ir a Nueva York en avión, la razón me dice que es mejor coger un avión que vaya a Nueva York que uno que vaya a Constantinopla⁹.

El mismo razonamiento vale para el caso de las aplicaciones de las teorías científicas. Los desarrollos de la física fueron utilizados para la fabricación de bombas nucleares porque muchas personas querían tener un método efectivo de ganar una guerra o de atemorizar a sus enemigos. Sin tales deseos, las bombas no hubieran existido. Sería absurdo, por otra parte, rastrear la intención de causar daño o provocar destrucción, directamente en la física de partículas. El asunto fue exactamente al revés: una vez descubierta la posibilidad de transformar pequeñas cantidades de materia en colosales despliegues de energía, alguien pensó en los usos pragmáticos que tal posibilidad ofrecía; pero como lo ilustra ricamente la ciencia ficción, posibilidades idénticas podrían redundar en nuestro beneficio en contextos bien diferentes. Algo similar ocurre con ciertos problemas sociales y políticos. Por ejemplo, gracias a ciertos avances científicos y tecnológicos, hoy es posible resolver el problema del hambre en el mundo. Pero para poner en marcha la solución se necesita primero el motor humeano: el deseo o la intención por parte de quienes tienen los medios. A este respecto, podríamos parafrasear al filósofo norteamericano Willard Quine, quien dijo, en una observación

⁹ RUSSELL, B. La sociedad humana: ética y política. Madrid: Cátedra, 1984. p. 10.

sobre epistemología¹⁰, que la situación humana sigue siendo todavía la situación *humeana*. Es evidente, por contraste, cuando dirigimos la mirada al campo de la ciencia, que sus características salientes parecen dar la razón a los partidarios de la tesis del neutralismo valorativo.

La actitud racional que está a la base de la ciencia puede explicarse apelando a dos componentes. El primero es el afán por encontrar explicaciones sistemáticas de los fenómenos naturales. En este sentido, la ciencia comparte una motivación fundamental con el mito y la religión: las tres empresas intentan hacer comprensible un mundo que se nos presenta extraño o cuando menos complejo. El otro componente es el que distancia a la ciencia de la religión y el mito: la crítica racional sistemática. Este segundo elemento es el que obliga a los científicos a presentar sus ideas de la forma más clara posible, con frecuencia en términos cuantificables. Es lo que explica la diferencia entre la mera profecía del palmista afortunado y la predicción científica; es lo que nos permite comprender por qué, mientras los científicos realizan experimentos para poner a prueba sus hipótesis, los predicadores y los profetas, en cambio, presentan sus sermones y advertencias como derivados de una fuente divina. La ciencia promueve la identificación del error y promueve el cuestionamiento severo de puntos de vista; la religión, por el contrario, exige la aceptación acrítica.

Si este análisis es correcto, entonces el mal no reside en la ciencia misma, concebida como una empresa racional cuyo fin primordial es la búsqueda de la verdad¹¹. Usando a Hume para nuestros propósitos, podríamos decir que, en tanto nos enfrentamos con un creador humano, el uso práctico de la ciencia es esclavo de las pasiones. Cuando tales pasiones son insanas, la ciencia se puede convertir en un instrumento para la destrucción, quizás el más poderoso y terrible. Cuando son adecuadas, la ciencia es otra vez la mejor herramienta que podemos emplear para el beneficio de todos.

No estamos diciendo, sin embargo, que el desarrollo de la ciencia esté motivado únicamente por la búsqueda desinteresada de la verdad. Muchas veces ocurre que una nueva teoría o hipótesis importante surge

¹⁰ QUINE, Willard V. O. "Epistemología naturalizada". En: *La Relatividad Ontológica*. Madrid: Tecnos, 1974. El comentario de Quine es acerca de las dudas de Hume sobre la confiabilidad de las inferencias inductivas.

¹¹ "Así pues, no son los científicos los que son moralmente ambiguos, sino la gente en general". SAGAN, Carl. *El mundo y sus demonios*. Bogotá: Planeta, 1997. p. 317.

porque alguien desea obtener algún beneficio, o porque vislumbra las posibilidades de conseguir alguna ventaja para el grupo social de sus afectos. Por ejemplo, la reacción renacentista contra la física de Aristóteles se vio favorecida por la necesidad de construir cañones cada vez más precisos. Uno de los contratistas, Tartaglia, al realizar las pruebas de sus cañones, se dio cuenta de que la trayectoria de las balas parecía ignorar las leyes aristotélicas del movimiento, cuya corrección daban por sentada los expertos. Sus observaciones sirvieron para dar al traste con las teorías físicas tradicionales, y aunque estaban motivadas por el afán bélico de sus patrocinadores, terminaron por contribuir al desarrollo de la ciencia. Más cerca de nosotros está el caso de las investigaciones en física nuclear financiadas con el propósito de sofisticar los mecanismos de destrucción del enemigo. Tales investigaciones resultaron en avances para la ciencia, y en una amenaza cada vez más terrible para la vida humana. Desde luego, como ya lo dijimos, los científicos que participaron en esos proyectos, *qua* individuos, tienen su cuota de responsabilidad moral.

Teniendo en mente los horrores que la ciencia y la tecnología han hecho posibles, algunos incluso han sugerido que deberíamos adoptar la conclusión radical conservadora de que, en definitiva, hay ciertas cosas que no deberíamos saber. Roger Shattuck, por ejemplo, ha defendido tal idea abiertamente. Haciendo una analogía con el conocido tema literario del Dr. Jekyll y Mr. Hyde, intenta persuadir a sus lectores de que la búsqueda de la verdad puede llegar a ser incompatible con el cultivo del talante moral. Veamos sus palabras: “El Dr. Jekyll (...) nos transporta hasta un ominoso dilema: la confrontación entre la verdad y sus consecuencias. Porque, aunque experimenta en su propia persona, *el obsesionado científico demuestra que la verdad puede tener consecuencias imprevisibles y devastadoras. Los evidentes peligros de sus experimentos le atraen en lugar de contenerle. El talento de Jekyll en la búsqueda de la verdad destroza su carácter moral*”¹².

Estas dudas, sin embargo, nos parecen infundadas. Las tesis de Shattuck sugieren que, a veces, el conocimiento puede ser la fuente del mal (de ahí el título de su libro: *Conocimiento Prohibido*). Pero, si Hume tenía razón, eso no puede ser cierto. Desde luego que es mucho más peligroso un asesino inteligente e ilustrado que uno bruto e ignorante. Pero, después de todo, los dos son igualmente asesinos y, por tanto, el

¹² SHATTUCK, Roger. *Conocimiento prohibido*. Madrid: Taurus, 1998. p. 19. Cursivas añadidas.

grado de sofisticación de sus conocimientos no puede ser la variable que haga la diferencia significativa. Kant había planteado un tema similar pero de forma distinta. Él se preguntó si podía haber algo que pudiera denominarse “bueno sin restricción”. Encontró que ninguno de los bienes naturales que tomamos como valiosos puede recibir esa denominación: “El entendimiento, el juicio, o como quieran llamarse los talentos del espíritu (...) son, sin duda, en muchos respectos, buenos y deseables; [pero] también pueden llegar a ser extraordinariamente malos y dañinos, si la voluntad que ha de hacer uso de estos dones (...) no es buena”¹³.

Aunque sabemos que las doctrinas morales de Hume y de Kant son profundamente inconsistentes entre sí, creemos que es posible combinar coherentemente los elementos que hemos mencionado de cada una de ellas, para ofrecer una explicación satisfactoria al problema de las relaciones entre la ciencia y el mal. Para tal cometido, resulta preciso insistir en la distinción que hemos introducido desde el comienzo de este trabajo sobre la naturaleza dual de los procesos a través de los cuales producimos la ciencia, y de aquellos mediante los que utilizamos sus realizaciones. Tal distinción impone una doble manera de entender y definir la ciencia misma. En primer lugar, la ciencia puede concebirse como una empresa puramente racional cuya motivación fundamental es la búsqueda de la verdad. Desde esta perspectiva, podría decirse que la ciencia es moralmente neutral, pero sólo porque la verdad es moralmente neutral también. En segundo lugar, la ciencia puede ser vista como una práctica humana en la que seguramente están involucrados los mismos intereses que están presentes en todas las prácticas humanas: afán de poder y reconocimiento, por ejemplo¹⁴. Desde esta perspectiva, la ciencia puede llegar a ser uno de los peores instrumentos del mal. Pero el problema, entonces, no está en la ciencia misma concebida como actividad racional, sino más bien en la naturaleza humana.

En este punto debemos regresar a la pregunta de si, dada la ancestral inclinación humana hacia la destrucción, no sería mejor obstaculizar el desarrollo de la ciencia, sobre todo en áreas como la física, la química y la biología, que han producido algunas de las armas más temibles. Nuestra respuesta es negativa. Pensamos que la actitud racional que la

¹³ KANT, Immanuel. *Fundamentación de la Metafísica de las Costumbres*. México: Porrúa, 1998. p. 21.

¹⁴ Que esta distinción es fundamental, lo comprueba el hecho de que ciertos filósofos contemporáneos han percibido la necesidad de llamar la atención de sus lectores sobre el papel que desempeñan los intereses individuales y colectivos en la búsqueda del conocimiento. Ver por ejemplo la popular propuesta de Habermas en su *Conocimiento e Interés*. Vide bibliografía.

ciencia requiere y promueve es un requisito de nuestra supervivencia como especie y, quizás también, de la búsqueda de soluciones a nuestros problemas más importantes.

Si, como pensaba Hume, la razón es esclava de las pasiones, entonces finalmente no hay un conflicto real entre ambos aspectos de la naturaleza humana. No lo hay porque la razón carece de fuerza motora. Como lo expresa Garrett Thomson en su comentario de las tesis morales de Hume:

Uno de los principios básicos de Hume es que el razonamiento, por sí solo, nunca puede motivar la acción. La razón es inerte por lo que respecta la motivación y, debido a esto, la razón nunca puede oponerse a una pasión en la dirección de la voluntad, a pesar del habla popular, e incluso filosófica, acerca del conflicto entre la razón y las pasiones (...) Sin un deseo o pasión relevante, somos completamente indiferentes hacia lo que es verdadero o falso¹⁵.

Este planteamiento tiene consecuencias radicales. Implica, por ejemplo, que un ser perfectamente racional también podría ser completamente cruel e inhumano. La racionalidad, según este punto de vista, se refiere no a los fines de nuestras acciones, sino meramente a los medios para alcanzar dichos fines. El propio Hume resalta pintorescamente estas implicaciones, cuando dice que “no es contrario a la razón preferir la destrucción del mundo entero a un rasguño en un dedo”¹⁶. Sin embargo, Hume también pensaba que el uso de la razón puede contribuir a modificar nuestros deseos o pasiones. Según él, que la razón esté subordinada a las pasiones no implica necesariamente que no haya pasiones irrazonables o erróneas. Por ejemplo, algunos de los deseos que tenemos pueden ser inadecuados porque se basan en información falsa. Uno puede tener el deseo irrefrenable de beber un vaso de agua que está envenenado, pero sólo porque no sabe que el vaso contiene veneno. De la misma manera, los miembros de una sociedad fanática pueden querer que ciertas personas desaparezcan del planeta, pero sólo porque aceptan alguna serie de proposiciones falsas acerca de dichas personas. Este es un tema común y aquí abundan los ejemplos. Muchas personas en las sociedades esclavistas creían firmemente que los negros no pertenecían a la misma especie que los blancos, que tenían

¹⁵ THOMSON, Garrett. *On Modern Philosophy*. Belmont, CA/US: Wadsworth Publishing Co Inc, 2004. p. 131.

¹⁶ HUME. *Op. cit.*

una dotación natural diferente que la del resto. Es muy probable que tales creencias apoyaran la convicción esclavista. Si es así, entonces la demostración de su falsedad pudo haber contribuido a la superación del racismo –si es que se puede hablar en estos términos. Las siguientes palabras de Carl Sagan resumen elocuentemente este punto de vista y, de hecho, parecen una reescritura de las tesis de Hume aplicadas ya no a la razón sino a la ciencia: “La ciencia por sí misma no puede apoyar determinadas acciones, pero sin duda puede iluminar las posibles consecuencias de acciones alternativas”¹⁷.

¿Cuál puede ser entonces la estrategia para enfrentar los problemas que plantea la enorme capacidad de destrucción que ofrece la ciencia cuando está al servicio de las pasiones equivocadas? Creemos que Hume puede, otra vez, decir cosas interesantes a este respecto. Para él, la moralidad se fundamenta en una clase especial de sentimientos que puede ser nombrada con el vago término ‘simpatía’. En ausencia de esta clase de emociones, la razón es impotente para introducir cualquier convicción moral o de consideración por los demás. O para decirlo de otra manera, la impasibilidad de la razón, si la consideramos totalmente aislada de las emociones, puede modificarse de modo favorable para los intereses de la especie humana, recurriendo a lo que muchos han identificado como la fuente del problema; a saber, la esfera de las emociones. Si los seres humanos albergamos emociones como la compasión sincera (la sensibilidad frente al sufrimiento de los demás) podremos abstenernos de actuar “irracionalmente” y privilegiar los usos pacíficos de los descubrimientos científicos por encima de los destructivos.

Para concluir, queremos reiterar la tesis cuasi humeana sobre las contribuciones que puede hacer la ciencia al progreso físico y moral de la humanidad. Tales contribuciones son posibles, pero sólo si el desarrollo de la ciencia se ve acompañado por el ejercicio de la compasión, o del conjunto de emociones al que Hume llamaba “simpatía”. En caso contrario, el progreso científico no hará más que favorecer la destrucción y el horror. Nuestros problemas más graves, entonces, no se podrán resolver con un aumento de nuestra racionalidad. Lo que debe cambiar es, más bien, la parte no racional de la naturaleza humana.

¹⁷ SAGAN. Op. cit., p. 45.

REFERENCIAS

- AYER, Alfred. (1952). *Language, Truth and Logic*. New York: Dover Books.
- FEYERABEND, Paul. (1987). *Ciencia en una Sociedad Libre*. Madrid: Siglo XXI.
- HABERMAS, Jürgen. (1989). *Conocimiento e Interés*. Madrid: Taurus.
- HUME, David. (1981). *Tratado de la Naturaleza Humana*. Madrid: Editora Nacional.
- KANT, Immanuel. (1988). *Fundamentación de la Metafísica de las Costumbres*. México: Porrúa.
- POPPER, Karl. (1996). *En busca de un mundo mejor*. Barcelona: Paidós.
- QUINE, Willard V. O. (1974). "Epistemología naturalizada". *En: La Relatividad Ontológica*. Barcelona: Tecnos.
- RUSSELL, Bertrand. (1984). *La sociedad humana: ética y política*. Madrid: Cátedra.
- _____. (1961). *The Basic Writings of Bertrand Russell*. New York: Simon and Schuster.
- SAGAN, Carl. (1997). *El mundo y sus demonios*. Bogotá: Planeta.
- SHATTUCK, Roger. (1998). *Conocimiento prohibido*. Madrid: Taurus.
- STEVENSON, C. L. (1959). "The emotive Meaning of Ethical Terms." *In: AYER, Alfred (Ed). Logical Positivism*. New York: The Free Press.
- SCHLICK, Moritz. (1959). "What is the Aim of Ethics?". *In: AYER, Alfred (Ed). Logical Positivism*. New York: The Free Press.
- THOMSON, Garrett. (2004). *On Modern Philosophy*. Belmont, CA/US: Wadsworth Publishing Co Inc.
- WEBER, Max. (1991). "Value Judgments in Social Science". *In: BOY, Richard et. al (Eds). The Philosophy of Science* Cambridge: The MIT Press.

PACTO SOCIAL Y LIBERTAD EN THOMAS HOBBS

SOCIAL PACT AND LIBERTY IN THOMAS HOBBS

JOHN FREDY LENIS CASTAÑO
Universidad de Antioquia, Colombia. johnlenisc@gmail.com

RECIBIDO EL 6 DE AGOSTO DE 2010 Y APROBADO EL 30 DE SEPTIEMBRE DE 2010

RESUMEN ABSTRACT

Este artículo pretende dilucidar algunas implicaciones antropológico-pedagógicas del pacto instituidor del Estado civil propio de la teoría política de Thomas Hobbes. Aquéllas tienen que ver, por un lado, con el sentido educativo de la instancia jurídico-penal y, por otro, con el tipo de libertad a la que pueden aspirar los ciudadanos en dicho Estado. Para ello se hará el siguiente recorrido: el pacto como instituidor del Estado civil; el sentido educativo del Estado; el concepto de libertad civil; y, finalmente, la ética del soberano y los límites del Estado.

This paper's aim is to dilucidate some of the anthropological-pedagogical implications of the pact that institutes the civil State, which is a concept that characterizes Thomas Hobbes' political theory. Such implications deal with, on the one hand, the educational sense of the law and its enforcement, and on the other, with the type of liberty that the citizens of the aforementioned State have. To achieve this, the following issues shall be discussed: the pact as a means for instituting the civil State, the educational sense of the State, the concept of civil liberty, and, finally, the ethics of the sovereign and the limits of the State.

PALABRAS CLAVE KEY WORDS

castigo, educación, Hobbes, libertad civil, límites del estado, pacto social.

punishment, education, Hobbes, civil liberty, limits of the State, social pact.

I. ENTRE EL ESTADO DE NATURALEZA Y EL ESTADO CIVIL, EL PACTO

En la explicación de la evolución social y política de los individuos, Hobbes recurre a una hipótesis que podría sintetizarse como la vía *estado de naturaleza-pacto-estado civil*. Un camino cuyo término medio constituye, precisamente, el punto de inflexión en el giro que va del hombre egoísta y autodestructor al ideal de ciudadano cabal que requiere toda convivencia social estable y segura; y con el cual se hace un corte temporal que instauro un antes y un después en la existencia humana, introduciendo un vital cambio de dirección en la naturaleza salvaje del hombre.

En el antes encontramos ese estado o “condición natural del género humano”¹ que consiste en la búsqueda ciega del provecho individual y que hace posible, en tanto condición universal, una situación de igualdad entre todos los seres humanos tanto por el poder que todos tienen de matar al otro, como por las mismas pasiones que les embargan². Características antropológicas que resumidas en el concepto de “egoísmo” devienen en agresión y lucha de todos contra todos cuando coinciden en el objeto deseado, dándose así una condición de miedo mutuo e inseguridad permanentes que ponen en cuestión cualquier concepto de “mío” o “tuyo” pues lo que uno le quita a otro le puede ser arrebatado, a su vez, por un tercero; y así sucesivamente en una guerra sin fin³. Este “estado de naturaleza” tiene como sentido principal en el marco general de la teoría hobbesiana la construcción de las bases mismas de su propuesta política, ya que con él, y por contraste, justificará la necesidad imperiosa que hay del Estado civil. En efecto desde el primer capítulo del *Ciudadano*⁴ y el decimotercero del *Leviatán*, esa urgencia aparece planteada de manera insistente, haciendo que el contraste entre ambos estados sea una especie de contradicción antitética. Veamos:

En una situación semejante no existe oportunidad para la industria, ya que su fruto es incierto; por consiguiente no hay cultivo de la tierra, ni navegación, ni uso de los artículos que pueden ser importados por mar, ni construcciones

¹ HOBBS, Thomas. *Leviatán*. Escohotado, Antonio (Tr). Madrid: Nacional, 1979. pp. 222ss.

² *Ibid.*, cap. VI, p. 156-168.

³ *Ibid.*, p. 223.

⁴ HOBBS, Thomas. *El ciudadano*. Madrid: C.S.I.C y Editorial Debate, 1993. p. 14-21.

confortables, ni instrumentos para remover las cosas que requieren mucha fuerza, ni conocimiento de la faz de la tierra, ni cómputo del tiempo, ni artes, ni letras, ni sociedad; y lo que es peor, existe continuo temor y peligro de muerte violenta; y la vida del hombre es solitaria, pobre, tosca, embrutecida y breve⁵.

O, tal como aparece en *El Ciudadano*:

Fácilmente se deduce lo poco idónea que es la guerra perpetua para la conservación tanto del género humano como de cada hombre.(...) Y todo el que pensase que habría que permanecer en aquel estado en el que todo es lícito para todos, se contradice a sí mismo. Ya que todo hombre, por necesidad natural, tiende a lo que es bueno para él, y no hay nadie que considere que esta guerra de todos contra todos, que es inherente por naturaleza a aquel estado, sea bueno para sí. Por eso sucede, por el miedo recíproco, que consideremos que hay que salir de semejante estado y conseguir aliados para que, si es que ha de haber guerra, que no sea de todos contra todos ni sin ayuda⁶.

Y aunque este concepto de “estado de naturaleza” es introducido por Hobbes como herramienta teórica o hipótesis basta que, como él mismo advierte, se haga una contrastación fáctica o empírica, para ver la actitud de alerta y prevención mutua que sostienen los seres humanos en la vida cotidiana, o la situación de algunos “grupos salvajes o primitivos”, o las características de la condiciones de vida de una sociedad cuando cae en una guerra civil, para constatar la necesidad del Estado civil⁷.

Sin embargo, y afortunadamente, en los seres humanos existe, al lado de las pasiones que producen la riña, otras pasiones que le hacen buscar la paz y, además, la razón necesaria para realizar los acuerdos conducentes a ella: “las pasiones que inclinan a los hombres hacia la paz son el temor a la muerte; el deseo de aquellas cosas que son necesarias para una vida confortable; y la esperanza de obtenerlas por su industria. Y la razón sugiere adecuados artículos de paz sobre los cuales puede llevarse a los hombres al acuerdo”⁸.

⁵ HOBBS, Thomas. *Leviatán o la materia, forma y poder de una república eclesiástica y civil*. México: Fondo de cultura económica, 1996. p. 103.

⁶ Op. Cit., p. 20.

⁷ Cfr. especialmente en el *Leviatán*, Op. cit., p. 225 y 226.

⁸ *Ibid.*, p. 227.

De esta manera se hace relevante el hecho de que si bien, en un primer momento, la naturaleza humana aparece únicamente como generadora de la guerra propia del “estado de naturaleza”; ahora, en un segundo momento y como transición al planteamiento de la constitución de una sociedad civil o Estado, esta misma naturaleza aparece como la posibilitadora, en principio, de esta conformación política. Algo que se convierte en un principio antropológico fundamental para la teoría y la práctica políticas puesto que si la naturaleza humana sólo encerrara el germen de la riña y la discordia y, por tanto, ningún ápice de ánimo para la conservación de la vida, y una vida confortable, la convivencia y la paz serían algo más que una utopía.

Ahora bien, la mera condición natural – razón y pasión connaturales a la humanidad que permiten el reconocimiento de la necesidad de ceder el propio poder en aras del bienestar personal – no es suficiente para realizar esa vida confortable que todos anhelan. Precisamente por ello se hace necesario el pacto y la espada estatal que surge con él. Al respecto Hobbes es claro cuando explica en el capítulo XVII del *Leviatán*⁹ que la realización de la finalidad de la constitución del Estado – que es conservar la vida y garantizar las condiciones necesarias para mantener la seguridad y la paz, haciendo de estas un contexto favorable para el logro individual de una vida dichosa – no es posible de esperar en una situación donde dicho pacto no exista, puesto que cuando falta un poder vinculante común que busque la paz, instituya el castigo, infunda el temor e instaure la obligatoriedad del cumplimiento de los convenios realizados, los contratos son meras palabras y las pasiones egoístas continúan señoreándose del actuar humano. Así, la generación de una república ocurre cuando cada uno de sus miembros decide como autor delegar en un hombre o asamblea de hombres su representación, con lo cual confiere todo su poder de autorregulación y fuerza a ésta, lo cual es un sometimiento de todas las voluntades a una y de todos los juicios a uno. Unidad de todos en uno que Hobbes define “(...)como si todo hombre debiera decir a todo hombre: autorizo y abandono el derecho a gobernarme a mí mismo, a este hombre, o a esta asamblea de hombres, con la condición de que tu abandones tu derecho a ello y autorices todas tus acciones de manera semejante”¹⁰.

⁹ *Ibid.*, p. 263-268.

¹⁰ *Ibid.*, p. 267.

De donde surge la definición de República: “(...) es una persona cuyos actos ha asumido como autora una gran multitud, por pactos mutuos de unos con otros, a los fines de que pueda usar la fuerza y los medios de todos ellos, según considere oportuno, para su paz y defensa común”¹¹.

Esa persona se convierte así en soberano y todos los demás en súbditos. Pero hasta aquí se ha hecho un recorrido meramente retrospectivo en tanto se ha mirado lo que hay antes del pacto y lo que se configura con él. Se hace necesario entonces, en este intento de análisis temporal, preguntar por la dimensión prospectiva, por el después de ese momento contractual. ¿Qué sucede entonces luego de la instauración del Estado civil: se da una metamorfosis total del ser humano, pasando de ser un individuo egoísta a ser un ciudadano magnánimo y ejemplar; o se continúa en una coexistencia marcada por la paradójica condición humana pasional-racional? Pensar lo primero sería tildar la teoría de Hobbes de ingenua. Optar por lo segundo es mucho más acorde con su realismo antropológico y político. Pero si el ser humano continúa siendo ese ser heterogéneo y la mayoría de las veces en pugna consigo mismo en el orden civil constituido, ¿cuál es el camino que, a partir de la obra de Hobbes, podría ser pensado como vía conducente a la realización gradual, pero cada vez mayor, del ideal de Estado propuesto por él? En aras de clarificar la posición misma de nuestro autor se podría comenzar recordando la crítica que él hace de la noción aristotélica de *zoon politikon*¹², y si de acuerdo a ésta es válido decir que el ser humano no nace apto para la sociedad civil sino que se hace, puede concluirse que el aspecto que mejor responde a esta última pregunta es la dimensión formativa del Estado.

II. EL SENTIDO EDUCATIVO DEL ESTADO Y DE SU ESPADA

La educación, en efecto, desempeña una función ético-política fundamental en el planteamiento filosófico de Thomas Hobbes teniendo en cuenta que ella se convierte en uno de los pilares sobre los cuales se constituye y se mantiene el Estado. Una función doble (ética y política) y simultánea que consiste básicamente en lograr una legitimación del Estado a través de la enseñanza, transmisión y adoctrinamiento en el corpus de las leyes civiles promulgadas. Y esto también por una exigencia lógico-jurídica: la ley tiene que ser divulgada y conocida

¹¹ *Ibidem*.

¹² Cfr. nota 1 del capítulo 1 del *Ciudadano*, HOBBS, Op. cit., p. 14 y 15

para poder obligar¹³. Ley que, además, “será más fácilmente entendida si está redactada sobriamente, ya que todas las palabras se prestan al equívoco (multitud de significaciones)”¹⁴, implicándose así – con este llamado a la simplicidad de la ley – una negación de que la exigencia o el criterio fundamental para la articulación del individuo en la sociedad civil sea la condición ilustrada de éste aunque, en su afán legitimador, la educación busque el mayor reconocimiento y aceptación racional posible de esas leyes por parte de los ciudadanos como reglas necesarias, no sólo como leyes a las cuales hay que obedecer ciegamente. Esto es, hacer ver de manera legítima y razonable la necesidad y el sentido del paso del “estado de naturaleza” al “Estado Civil”. Algo que muestra, de suyo, la conciencia tan grande que tenía el filósofo inglés de la importancia de las propias convicciones en el mantenimiento del orden social y político. Él sabía que la enseñanza tenía que ser una prioridad en el conjunto de las tareas estatales si se quería intentar, de una manera racional¹⁵ y responsable, llegar a la conciencia de cada ciudadano y evitar así un desmoronamiento del Estado causado por la incredulidad y el desconocimiento de las causas de la guerra¹⁶ propios de la condición de ignorancia de los individuos cuando no ha existido un proceso formativo adecuado del pensamiento y de la capacidad de calcular (razonar). Por ello es que en el capítulo XII de *El Ciudadano* y en el capítulo XXIX del *Leviatán*¹⁷ considera como una de las causas de la disolución del Estado la enseñanza de falsas doctrinas¹⁸.

Pero la educación no viene a significar – como bien lo ha hecho ver Schmitt – una posibilidad de perfeccionamiento moral que permite ingresar de una vez y para siempre en el reino armonioso donde todos, en tanto sujetos absolutamente morales y magnánimos, actúan por puro respeto a la dignidad humana, “Hobbes está muy lejos de tales concepciones. También en su teoría encontramos la posibilidad de influir

¹³ HOBBS, *Leviatán*, Op. cit., p. 338-368.

¹⁴ TÖNNIES, Ferdinand. *Hobbes, vida y doctrina*. España: Alianza, 1988. p. 299

¹⁵ Recordemos que en el *Tratado sobre el cuerpo* Hobbes define enseñar como demostrar y en este sentido como una actividad plenamente racional. Cfr. en la edición de Trotta (Madrid, 2000, p. 85).

¹⁶ Al respecto es aclaratoria la observación de F. Tönnies cuando dice: “[Hobbes] También espera un progreso cultural por mediación de la Ciencia, pero en especial de aquella parte de la misma que se cree fundador, el *Derecho natural*, base común de la Ética y de la Política. Tiene por objeto acabar con las luchas y guerras, evitar las civiles, ‘pues su causa está en nuestro desconocimiento de las causas de la guerra y de la paz’; y de ellas proceden la miseria y todo lo que es aborrecible para el hombre.” Op. cit, p. 231.

¹⁷ Op. cit, p. 103-111 y 395-406, respectivamente.

¹⁸ Cfr. igualmente sobre la importancia de la educación en la teoría de Hobbes: BOBBIO, N. *Thomas Hobbes*. México: F.C.E., 1992. p. 38 y OAKESHOTT, M. *El racionalismo en la política y otros ensayos*. México: F.C.E., 2000. p. 248.

en los hombres por la coacción y la educación, pero él no se hace grandes ilusiones sobre la naturaleza humana. Él ve al hombre muy 'insocial', como un animal, lleno de miedo y preocupación ante el futuro,(...)"¹⁹.

Pretensión de influencia racional y pasional — por el miedo al castigo — que, sin embargo, en ningún caso tiene el sentido de una fusión del fuero externo con el fuero interno (asunto que será reflexionado en el cuarto y último numeral de este ensayo). Por otra parte, la educación también funda su posibilidad de ejercicio legitimador en el hecho de que es una actividad mediada por el lenguaje. Recordemos, como lo plantea Zarka²⁰, que es la actividad lingüística la que permite al ser humano comunicarse con sus congéneres y razonar y acordar convenios de paz y provecho mutuos. La enseñanza de la necesidad, el funcionamiento y el contenido de esos pactos es un imperativo inomitible para cualquier Estado que pretenda fundarse racionalmente. Pero además, y tal como aparece elaborado por Hobbes²¹ y resaltado por Zarka²², el lenguaje también encierra en sí mismo la potencialidad de engañar, algo que alerta sobre la necesidad de no confiar ni en la mera racionalidad ni en el mero acto de prometer; es esta posibilidad de mentir y de incumplir lo pactado lo que impone el principio de educar también a través del castigo. Educación punitiva que, a través de amenazas penales y condenas efectivas, busca disciplinar las pasiones enemigas del Estado o, en palabras de Oakeshott, lograr la "moralidad del hombre domado"²³. Lo cual no significa que Hobbes desconociera que, en medio de los miles de seres humanos que conforman los distintos grupos, existen seres "magnánimos" no necesitados de esos artificios jurídicos para obrar moral y políticamente de manera correcta. Hobbes sabía que estos caracteres existían. El punto es que él estaba interesado en resolver el problema de qué hacer con aquellos que no tenían tal temperamento, pues el dilema se debía, precisamente, a "la escasez de caracteres nobles"²⁴. Y la respuesta que este filósofo da es: castigar. Un consejo que no se debe a algún tipo de sadismo penal que pudiera venir a justificar la exigencia, desde alguna perspectiva no realista, de la desaparición de los castigos fuertes. No, la espada y la educación en la teoría de Hobbes se articulan en una tensión permanente entre el deseo de contravenir la ley por el impulso

¹⁹ SCHMITT, C. *El Leviatán en la teoría del Estado de Thomas Hobbes*. México: Universidad Autónoma Metropolitana, 1997. p. 81.

²⁰ ZARKA, Yves Ch. *Hobbes y el pensamiento político moderno*. Barcelona: Herder, 1997. p. 95 y 96.

²¹ *Leviatán*, Op. cit., cap. 17.

²² Op. cit., p. 150 y 151.

²³ Op. cit., p. 315.

²⁴ *Ibid.*, p. 318

egoísta y el miedo frente a la amenaza del castigo; entre el espíritu del león y el del buey; entre el estar dispuesto a arriesgar la vida en aras del cumplimiento de un poderoso deseo y el mismo instinto de autoconservación²⁵. Asignándose de esta manera un sentido positivo y necesario al castigo. Al respecto podemos decir de la mano del mismo Hobbes que “una pena es un mal infligido por la autoridad pública a quién ha hecho u omitido lo que esa misma autoridad considera una transgresión de la ley. A fin de que la voluntad de los hombres esté por ello mismo mejor dispuesta a la obediencia”²⁶, “porque el fin del castigo no es obligar la voluntad del hombre sino formarla, y hacerla tal como la quiso el que instituyó el castigo”²⁷.

Finalidad entonces que permite que la pena pueda aspirar, no sólo a dar la posibilidad de un futuro al condenado, en tanto éste salga mejor dispuesto socialmente, sino también a ser considerada un mecanismo preventivo que ayude a los actores sociales “inocentes” a mantenerse a raya en esa tensión antes descrita. Pero ¿qué significa “hombre domado”, significa hombre sumiso y esclavo, hombre enajenado de todas sus potencialidades? Es en este punto donde la cuestión del arquetipo antropológico a formar implicado por este concepto hobbesiano de educación se revela como aspecto urgente para el análisis. Pasemos entonces al tercer numeral de esta indagación.

III. LIBERTAD CIVIL: SUMISIÓN CIEGA O POSIBILIDAD DE REALIZACIÓN

En todo asunto educativo hay un problema antropológico inherente: el tipo de hombre que se quiere lograr a través del proceso formativo. Modelo de hombre que según la doctrina hobbesiana debe tener como

²⁵ Punto donde —a diferencia de la crítica husserliana a este principio de autoconservación consistente en afirmar que Hobbes hizo del egoísmo un principio antropológico universal y único, con lo cual desconoció la “*das volle Menschenwesen*” (esencia total humana) (SAN MARTÍN, Javier. “Ética, antropología y filosofía de la historia. Lecciones de Husserl de Introducción a la Ética del Semestre de Verano de 1920”. En: Isegoría, Mayo de 1992. no. 5, p. 46-49)— hay que anotar que la atención del filósofo británico se centró en la búsqueda de un principio fundamental y, si se quiere, esencial, que permitiera la explicación del acceso de todos los individuos a la conformación del Estado civil; lo cual no quiere decir, como ya ha sido apuntado más arriba al hacer referencia a la escasez de caracteres nobles o magnánimos, que en su planteamiento se desconociera que en la condición humana existe un complejo de rasgos que no se reducen al egoísmo y a la búsqueda de autoconservación. En este sentido la crítica que etiqueta la filosofía hobbesiana de reduccionista es inadecuada y, paradójicamente, reductora.

²⁶ HOBBS, *Leviatán*, Op. cit, cap. XXVIII, p. 386. Cursivas en el original; el subrayado es mío.

²⁷ HOBBS, *El ciudadano*, Op. cit, p. 120.

principales rasgos la capacidad para vivir en sociedad y la capacidad para obedecer los mandatos del soberano. Esto es, una educación fundamentalmente cívica o ciudadana²⁸ que tiene la ventaja de introducir al individuo “egoísta” en la consideración de la alteridad como *conditio sine qua non* de su propia realización personal y que conlleva una reestructuración de la propia libertad. Recordemos que en el capítulo XXI del *Leviatán* la libertad es definida como la inexistencia de impedimento al movimiento y en este sentido el hombre libre “es quien en las cosas que por su fuerza o ingenio puede hacer no se ve estorbado en realizar su voluntad”²⁹.

Esta concepción niega el *libre albedrío* como una mera libertad del querer (léase voluntad, deseo o inclinación) y relaciona a aquél con el miedo y la necesidad. Con el primero en tanto lo que se hace por miedo también pudo haberse omitido y con la segunda en tanto que en el mundo nada sucede por azar, todo está condicionado por un orden causal, incluyendo no sólo los actos o acciones sino también los deseos e inclinaciones. Tratándose hasta aquí de una libertad meramente natural o, si se quiere, física al modo de una posibilidad de movimiento que habla de un primer nivel de la libertad humana y que, siendo el antes de la libertad civil —o de la libertad instaurada por el pacto—, refleja el condicionamiento fundamental de la situación humana, a saber, el condicionamiento natural que hace a los hombres, en tanto organismos vivos, depender de la propia estructura física, la energía disponible, etcétera, para existir y moverse en pos de los objetos deseados y tratar de obtenerlos en esa búsqueda incesante de felicidad que no es más, según Hobbes, que la satisfacción de los deseos cambiantes que impulsan en ese movimiento sin fin propio del sujeto eternamente deseante que es el ser humano. Tengamos en cuenta que este autor fue uno de los primeros filósofos, influenciados por los descubrimientos y avances de la ciencia natural moderna en definir al ser humano como otro ser más de la naturaleza y por tanto determinado por el orden causal de los fenómenos³⁰. Precisamente en su *Tratado sobre el cuerpo* hace una explicación exhaustiva de los conceptos de causa y de movimiento³¹; que lo coloca, propiamente, como uno de los representantes del mecanicismo moderno.

²⁸ Importancia que también es manifestada en “los deberes del soberano” planteados en el capítulo XIII de *El Ciudadano* y en el capítulo XXX del *Leviatán*.

²⁹ *Ibid.*, p. 300.

³⁰ Cfr. por ejemplo OAKESHOTT, Op. cit., p. 225; y TÖNNIES, Op. cit., p. 200, 206.

³¹ Op. cit., cfr. por ejemplo, p. 76-80.

De otra parte hay, en el marco estatal y civil, un conjunto de cadenas artificiales, las leyes, cuya existencia hace que se tenga que hablar de *la libertad de los súbditos* o, en nuestros términos, de libertad civil. Esta consiste en las cosas que el soberano ha omitido al legislar; algo que no niega el poder ilimitado del soberano porque sigue siendo cierto en la teoría de Hobbes que, según la tesis del autor/actor, el gobernante no puede ser acusado de injusticia por las determinaciones que tome. Punto a partir del cual el filósofo inglés quiere precisar en qué consiste la verdadera libertad de los súbditos, ya que ve en la tradición que viene desde la antigua Grecia el malentendido que confunde la libertad individual con la libertad de las Repúblicas. Algo para lo cual va a reflexionar sobre las cosas que un súbdito puede negarse a hacer, sin cometer injusticia, aunque hayan sido ordenadas por el soberano, esto es, “aquellas cosas cuyo derecho no pueden transferirse por pacto”. Así pues tenemos que el súbdito tiene derecho a desobedecer cuando se le exige que atente contra su propio cuerpo o contra la vida de otro ciudadano; cuando se le exige que se acuse a sí mismo; cuando se le exige que, venciendo el miedo, vaya a la guerra (donde hay que diferenciar la cobardía de la injusticia); cuando se le exige que no use las armas para su propia defensa; cuando se le dice que tiene que hacer algo que no está contemplado en el marco jurídico; cuando se le demanda que no devenga súbdito de un soberano que le ha vencido y tomado preso, ofreciéndole luego la seguridad y libertad que cualquier ser humano busca; cuando por falta de nombramiento de sucesor — por renuncia o por muerte sin haber hijos o sin haber hecho una declaración expresa sobre el sucesor — el puesto de soberano queda vacío; cuando el súbdito es desterrado y queda sometido a las órdenes del soberano de su nuevo país; cuando el soberano se declara vencido por otro y se hace su súbdito. Libertad que, en suma, podemos sintetizar con Hobbes en el principio que reza: “*la obligación de los súbditos para con el soberano se sobreentiende que dura tanto como el poder mediante el cual éste es capaz de protegerlos*”³². Y que, como ya fue dicho arriba, en un Estado civil constituido y eficiente, se aúna a ese otro principio del silencio de la ley. Algo a partir de lo cual, y teniendo en cuenta los causales que legitiman la resistencia de los súbditos, se puede responder a la pregunta sobre el significado de la expresión “hombre domado” diciendo, con palabras de Cassirer, que “no hay ningún *pactum subjectionis*, ningún acto de sumisión por el cual el hombre pueda renunciar a su condición de agente libre y esclavizarse a sí mismo pues, con semejante acto de renuncia, el hombre perdería

³² *Ibid.*, p. 309.

precisamente ese carácter que constituye su naturaleza y su esencia: perdería su humanidad”³³. Pérdida que vendría a significar la anulación de la posibilidad de construcción y realización de la personalidad, lo cual implicaría una contradicción con la finalidad misma del estado civil promulgado por Hobbes: garantizar las condiciones y la paz necesarias para la realización de una vida dichosa³⁴. En este sentido puede decirse que la sumisión del ciudadano es directamente proporcional a la capacidad del Estado para efectuar su cometido y, por esta razón, no se puede hablar en Hobbes de un absolutismo político³⁵. En esta línea, “‘libertad’ quiere decir, que la libertad natural no será limitada más de lo que exija el bien común, y que las gentes de buena voluntad no vayan a caer en la ley como en un lazo”³⁶.

Finalmente la libertad civil deviene, en el planteamiento que nos ocupa, en la condición de realización de una vida humana digna. Vida que no podría ser efectuada en un estado donde imperase la libertad natural, como en el estado de naturaleza, porque, de una parte, allí el derecho a todo termina siendo un derecho a nada y, de otra, el reconocimiento³⁷ y la relación interpersonal respetuosa, no existen en esa difícil situación natural.

Ahora bien, plantear la proporcionalidad entre sumisión ciudadana y eficacia estatal implica revisar la existencia de una ética del soberano y de unos límites inherentes al Estado.

IV. A MANERA DE CONCLUSIÓN: LA ÉTICA DEL SOBERANO Y LOS LÍMITES DEL ESTADO

Con lo dicho sobre la proporcionalidad que debe existir entre sumisión ciudadana y eficacia estatal, el concepto hobbesiano del Estado empieza a mostrar de manera explícita su carácter ideal. Para empezar a explicar esta afirmación puede apelarse a la idea de “ética del soberano”. Una

³³ CASSIRER, Ernst. El mito del estado. Santafé de Bogotá: F.C.E., 1996. p. 208. Al respecto Cfr. también OAKESHOTT, Op. cit., p. 264 donde acertadamente dice: “Ni es el efecto generado, el Leviatán, una destrucción deliberada del individuo; es, en efecto, la condición *mínima* de toda asociación entre individuos.”

³⁴ Cfr. OAKESHOTT, Op. cit., p. 250.

³⁵ Cfr. también ZARKA, Op. cit., p. 243 y 250

³⁶ TÖNNIES, Op. cit., p. 263.

³⁷ OAKESHOTT, Op. cit., p. 241.

noción que ha sido bien tematizada por varios estudiosos de la filosofía hobbesiana y que puede resumirse, por ejemplo, en las palabras de Zarka cuando dice:

El respeto a las leyes de naturaleza y el conocimiento del arte de gobernar definen lo que podemos llamar una ética del soberano, que consiste en un pleno ejercicio de la soberanía con todos sus derechos, pero también con sus deberes. Tomando libremente un concepto de Weber, podría caracterizarse esta ética del soberano como una ética de la responsabilidad. Ésta en Hobbes contiene una dimensión moral y una dimensión política, porque envuelve la consideración de los efectos o de las consecuencias de un acto. Así, instruir al pueblo sobre el fundamento de la ley, dictar buenas leyes (es decir, leyes necesarias para el bien del pueblo), no cometer actos de hostilidad contra los súbditos, son, entre otros, los caracteres esenciales de la ética del soberano como ética de la responsabilidad³⁸.

Y es precisamente en el derecho natural y todos aquellos causales de resistencia legítima donde el Estado hobbesiano halla sus propios límites. Por eso no se trata de un poder absolutista; se trata de un poder usado sabiamente por el soberano. Residiendo también, en esta pretensión del uso sabio del poder y la fuerza, su carácter de Estado ideal. Para Hobbes la legitimidad del Estado se funda en su propia perfección o eficacia, tanto en el ejercicio de su tarea protectora y garantizadora de paz como en su legislación y correcta aplicación de penas y castigos³⁹.

Es propio de un poder estatal racional de este tipo (...) la responsabilidad de la protección y seguridad de los sometidos al Estado. Al cesar la protección, también cesa el Estado y cae todo deber de obediencia. Entonces el individuo reconquista su libertad 'natural'. La relación

³⁸ ZARKA, Op. cit., p. 269. Cfr. también a este respecto STRAUSS, ¿Qué es filosofía política? Madrid: Guadarrama, 1970. p. 260 y SCHMITT, Op. cit., pp.77, 78 y 137.

³⁹ Con respecto a la legitimación del derecho a castigar Zarka concluye de manera atinada que, ante la antinomia entre el derecho a la resistencia de los súbditos y el derecho a castigar del soberano y su imposible resolución *a priori*, lo que queda es la legitimación *a posteriori* que se da cuando, al ser efectuado el castigo, los súbditos están convencidos de la justeza y no hostilidad de la actitud punitiva del soberano. Op. Cit., p. 268-270. ("La antinomia se formula así: si el derecho de resistencia es inalienable, los súbditos nunca han concedido al soberano el derecho a castigarlos, derecho que no puede concebirse como un atributo esencial de la soberanía, procedente del pacto que instituye el Estado. Inversamente, si el derecho penal es un atributo inalienable de la soberanía, fundado en el pacto social, el derecho de resistencia no puede ser considerado como un derecho inalienable del hombre" (p. 254)).

‘protección-obediencia’ es la piedra angular de la construcción hobbessiana del Estado⁴⁰.

En este sentido y con esta última afirmación tomada de Schmitt se abre una problemática que, en la perspectiva de este ensayo, se hace fundamental, a saber: ¿qué pueden hacer los súbditos adecuadamente educados en las leyes del Estado —esto es, con una capacidad bien desarrollada para calcular el grado de hostilidad de las acciones llevadas a cabo por el soberano— cuando éste no realiza correctamente sus funciones? Plantear una posibilidad de oposición crítica a la manera de la teoría crítica de la educación —con su enarbolado imperativo de formar el juicio crítico de los sujetos tal y como ha sido elaborado a partir de la escuela de Frankfurt, especialmente a partir de los planteamientos de Theodor Adorno y Max Horkheimer⁴¹— que coincidiera y se articulara con lo desarrollado por Hobbes sería una extrapolación riesgosa. La negación hobbessiana de la oposición y la desobediencia frente al Estado arguyendo que éstas son ilegítimas e irregulares, niega la posibilidad de que la educación sea una actividad fomentadora del sentido crítico aplicado —bajo la forma, por ejemplo, de la desobediencia civil— si tenemos en cuenta que la educación en esta concepción es un apéndice que está al pleno servicio del poder absoluto del Estado. Educación entonces que sólo debe propender por la supervivencia del soberano y del Estado instituido con él, quedando excluida toda posibilidad de cuestionamiento de éstos en el proceso de información y transmisión de las leyes de una generación a otra; algo que podría ser interpretado como una educación que busca hacer de los seres humanos meras piezas del engranaje de la máquina estatal en la búsqueda de ajuste fluido de éstas al entramado social tal y como el Estado le concibe. Pero ello es así porque, como ya ha sido anotado, para nuestro filósofo, cuando se habla de Estado civil, se presupone el buen funcionamiento de éste. Es decir, o estamos en un Estado/idea cuyo soberano y cuyos súbditos cumplen cabalmente sus roles —mandar respetando el derecho natural y obedecer acatando las leyes civiles, respectivamente— o estamos, otra vez, en el estado de naturaleza, el caos y la inseguridad. Mi hipótesis aquí es que, además de esta concepción rigurosa del Estado que impera en el pensamiento político de nuestro autor, tal disyuntiva se debe

⁴⁰ SCHMITT, Op. cit., p. 137.

⁴¹ Cfr. ADORNO, T. *Prismas*. Barcelona: Ariel, 1962. p. 9-29; *Mínima Moralia*. Madrid: Taurus, 1987. p. 41-42; *Consignas*. Buenos Aires: Amorrortu, 1973. p. 80-96. HORKHEIMER, M. *Teoría Crítica*. Barcelona: Barral, 1973. 229p. (especialmente el capítulo titulado “Razón y autoconservación”, p. 141-176), y ADORNO, T. y HORKHEIMER, M. *Dialéctica del Iluminismo*. Buenos Aires: Sur, 1970. p. 146-200.

también al concepto de verdad que existe en su teoría epistemológica. Recordemos que en su *Tratado sobre el cuerpo* trae una concepción del conocimiento que pone la verdad⁴², no sólo del lado de las verdades lógicas irrefutables (por ejemplo las que en su tiempo eran consideradas como tales, a saber, las de la geometría y las matemáticas), sino también haciendo parte de lo que debe ser aprendido cabalmente por parte de los educandos, si ella ha sido enseñada de manera adecuada (partiendo de las definiciones, logrando un conocimiento causal de los fenómenos, etc.). En este sentido y si el alumno aprende de forma clara y lógica lo que se le enseña en materia política, su actitud consecuente debe ser asumir sin discusión lo que, de suyo, es claro y evidente cuando ello – en este caso las leyes civiles – es demostrado racionalmente.

De este modo la teoría de Hobbes ofrece un planteamiento que pone como ideal cívico-político el equilibrio entre verdad-obediencia y mandato; entre el acto de prometer y las condiciones de realización de la promesa; entre la convicción propia del fuero interno (o legitimación del poder político en la conciencia de los ciudadanos) y la legislación y penalización propias del fuero externo y del uso del poder estatal. Asunto, este último, que nos lleva a la clarificación anunciada en el numeral dos y que concierne a la relación entre fuero externo y fuero interno. Al respecto, y aunque Schmitt plantea la tesis de que es precisamente dejar el fuero interno en el ámbito de lo que constituye el núcleo subjetivo e intocable del ciudadano por parte del poder del Estado lo que siembra la semilla de la destrucción del gran Leviatán⁴³; puede argüirse que esta separación hace de la propuesta hobbesiana algo todavía más realista. Esta separación no obedece a una miopía argumentativa existente en esta propuesta teórica; lo que ella hace es reconocer los límites de la coacción estatal⁴⁴ puesto que el único acceso que el Estado puede ambicionar en ese núcleo subjetivo es el de la demostración y el convencimiento racional, no sólo de la necesidad del orden político sino también de la justeza de sus leyes. Y ello sólo puede ser llevado a cabo en un proceso formativo que, al final, aparece como el sostenedor de una relación necesaria y vital para la existencia del mismo Estado: el buen desempeño del soberano y su reconocimiento por parte de los ciudadanos. En esta línea es ilustrativo el llamado que Hemón hace a su padre (Creonte) frente a su decisión de castigar a Antígona:

⁴² Op. cit., cfr. especialmente las pp. 87,88 y 89.

⁴³ SCHMITT, Op. cit., cfr. por ejemplo, p. 61, 111 y 116.

⁴⁴ Cfr. también STRAUSS, Op. cit., p. 262-264.

Padre, el más sublime don que de todas cuantas riquezas existen dan los dioses al hombre es la prudencia (o sabiduría). Yo no podría ni sabría explicar por qué tus razones no son del todo rectas; sin embargo, podría una interpretación en otro sentido ser correcta. Tú no has podido constatar lo que por Tebas se dice; lo que se hace o se reprocha. Tu rostro impone respeto al hombre de la calle; sobre todo si ha de dirigirse con palabras que no te daría gusto escuchar. A mí, en cambio me es posible oírlos, en la sombra, y son: que la ciudad se lamenta por la suerte de esta joven que muere de mala muerte, como la más innoble de todas las mujeres, por obras que ha cumplido bien gloriosas⁴⁵.

Puede concluirse entonces que la teoría de Hobbes no consiste en un recetario político orientador sobre qué hacer, de manera exacta y precisa en cada situación particular para evitar el conflicto en la vida humana, tanto de cara a las funciones del soberano como de cara a los deberes de los ciudadanos. Su proyecto es legitimador y fundamentador y contiene, como principio esencial y como ideal político, el imperativo de un obrar justo, esto es, un obrar que no permita decir que “la ciudad se lamenta” por el proceder del soberano o que el ciudadano actúa en contra de los fines mismos de la sociedad civil.

REFERENCIAS

- ADORNO, T. (1962). *Prismas*. Barcelona: Ariel.
- _____. (1973). *Consignas*. Buenos Aires: Amorrortu.
- _____. (1987). *Minima Moralia*. Madrid: Taurus.
- ADORNO, T. y HORKHEIMER, M. (1970). *Dialéctica del Iluminismo*. Buenos Aires: Sur.
- BOBBIO, N. (1992). *Thomas Hobbes*. México: Fondo de Cultura Económica.
- CASSIRER, Ernst. (1996). *El mito del estado*. Santafé de Bogotá: Fondo de Cultura Económica.
- HOBBS, Thomas. (1979). *Leviatán*. Escohotado, A. (Trad). Madrid: Nacional.
- _____. (1993). *El ciudadano*. Madrid: C.S.I.C. y Editorial Debate.
- _____. (1996). *Leviatán o la materia, forma y poder de una república*

⁴⁵ SÓFOCLES. *Antígona*. Navarra: Salvat, 1970. p. 99.

- eclesiástica y civil. México: Fondo de Cultura Económica.
- HORKHEIMER, M. (1973). *Teoría Crítica*. Barcelona: Barral.
- OAKESHOTT, M. (2000). *El racionalismo en la política y otros ensayos*. México: Fondo de Cultura Económica.
- SAN MARTÍN, Javier. (Mayo de 1992). "Ética, antropología y filosofía de la historia. Las lecciones de Husserl de Introducción a la ética del Semestre de Verano de 1920". *En*: *Isegoría*, no. 5, p. 46-49.
- SCHMITT, C. (1997). *El leviatán en la teoría del Estado de Thomas Hobbes*. México: Universidad Autónoma Metropolitana.
- SÓFOCLES. (1970). *Antígona*. Navarra: Salvat.
- STRAUSS. (1970). *¿Qué es filosofía política?* Madrid: Guadarrama.
- TÖNNIES, Ferdinand. (1988). *Hobbes, vida y doctrina*. España: Alianza.
- ZARKA, Yves Ch. (1997). *Hobbes y el pensamiento político moderno*. Barcelona: Herder.

LA TEMPOREIDAD EXTÁTICO-HORIZONTAL COMO ORIGEN DE LA TRASCENDENCIA DEL DASEIN

ECSTATIC-HORIZONTAL TEMPORALITY AS THE ORIGIN OF DASEIN'S TRASCENDENCY

CARLOS DI SILVESTRE

Universidad Nacional de Cuyo, Argentina. carlosds@logos.uncu.edu.ar

RECIBIDO EL 17 DE AGOSTO DE 2010 Y APROBADO EL 30 DE SEPTIEMBRE DE 2010

RESUMEN ABSTRACT

El artículo examina un aspecto relevante de la interpretación heideggeriana del tiempo como temporeidad del Dasein: su determinación como *horizonte trascendental* de la comprensión del ser en general. Si bien es cierto que Heidegger abandona el proyecto ontológico, diseñado en *Ser y tiempo*, de una interpretación *temporaria* del sentido del ser, el artículo intenta mostrar que, en textos posteriores, mantiene y profundiza su concepción de la temporeidad como horizonte trascendental de comprensión. Para ello, ofrece una reconstrucción de la noción de *esquema horizontal* de los éxtasis tempóreos, a fin de aclarar en qué sentido el tiempo constituye el origen ontológico del ser-en-el-mundo (trascendencia) como ámbito abierto de manifestación significativa.

This article examines a eminent aspect of the Heideggerian interpretation of the time as temporality of Dasein: its determination as *transcendental horizon* of the comprehension of Being in general. In spite of the fact that Heidegger abandons the on *Being and Time* designed ontological project of a *temporal* interpretation of the meaning of Being, the article tries to show that in posterior texts maintains and deepens his conception of the temporality as transcendental horizon of comprehension. In order to do so, it offers a reconstruction of notion *horizontal schema* of the temporal Ecstases, in order to explain how the time is the ontological origin of the being-in-the-world (transcendence) as open space for the sense apparition.

PALABRAS CLAVE KEY WORDS

esquema horizontal, Heidegger, ser-en-el-mundo, temporeidad, trascendencia.

being-in-the-world, Heidegger, horizontal schema, temporality, transcendence.

El propósito fundamental del proyecto teórico contenido en *Ser y tiempo* es, como se sabe, la elaboración concreta de la pregunta que interroga por el sentido del ser. Para ello, Heidegger se propone como meta *provisional* llevar a cabo una interpretación del tiempo como *horizonte de posibilidad* de toda comprensión del ser. Pero en la parte publicada de la obra, más exactamente, en su segunda sección, titulada "Dasein y temporeidad", la interpretación heideggeriana no llega aún al punto donde se habría alcanzado la mencionada meta provisional. En efecto, la segunda sección contiene ciertamente una interpretación del tiempo como temporeidad (*Zeitlichkeit*) del Dasein, es decir, como fundamento ontológico de la estructura del cuidado (*Sorge*) y, en cuanto tal, como tiempo originario a partir del cual se deriva la comprensión usual, "vulgar", y filosófica del tiempo como tiempo-ahora. Sin embargo, estos desarrollos no contienen aún la anunciada interpretación del tiempo como *horizonte trascendental* de la comprensión del ser en general, temática que debió ser expuesta en la tercera sección del tratado, titulada "Tiempo y ser"¹. El mismo Heidegger aclara posteriormente que el plan diseñado en SZ no pudo ser cumplido a causa de dificultades insalvables, vinculadas fundamentalmente con la pervivencia del lenguaje de la metafísica en la Ontología fundamental². Estas dificultades lo condujeron finalmente al abandono del proyecto ontológico de hacer comprensible el sentido del ser a través de una interpretación temporaria (*temporale*) del ser en cuanto tal, interpretación que debía transformarse a su vez en el hilo conductor para la destrucción de la historia de la ontología.

Sin embargo, el abandono del mencionado proyecto ontológico-histórico no implica que Heidegger haya abandonado también su concepción de la temporeidad como horizonte trascendental de la comprensión del ser. Antes bien, esta concepción de la temporeidad se mantiene posteriormente y es explicitada de un modo sistemáticamente consistente. Ella supone que el ser-en-el-mundo, en cuanto estructura de *trascendencia*, es un fenómeno que se constituye tempóreamente, es decir, que se origina en la temporización *extático-horizontal* del ser del Dasein. El presente estudio se propone aclarar este punto, apenas esbozado en SZ, y que, por ende, requiere de un examen de textos heideggerianos del período inmediatamente posterior al *Hauptwerk* de 1927. A partir de ellos explicaremos, en primer término, el sentido de la caracterización del ser-en-el-mundo como trascendencia. Luego,

¹ Cfr. HEIDEGGER, Martin. *Sein und Zeit*, 17a. ed., Tübingen, 1993. §§ 5, 8. En adelante SZ.

² Cfr. HEIDEGGER, Martin. "Brief über den Humanismus". In: *Wegmarken*, Frankfurt a. M. 1976, p. 328.

realizaremos una reconstrucción de la noción de *esquema horizontal* de la temporeidad y, finalmente, mostraremos cómo esta noción provee la base explicativa para la interpretación de la temporeidad como origen de la trascendencia.

1. EL SER-EN-EL-MUNDO COMO ESTRUCTURA DE TRASCENDENCIA

La noción de trascendencia ocupa un lugar relativamente marginal en SZ³. Sin embargo, luego va ganando en importancia hasta culminar en el escrito *Vom Wesen des Grundes*⁴ de 1929, donde se convierte en uno de los ejes centrales de la reflexión. Por ello, en lo que sigue nos ocuparemos fundamentalmente de este texto para mostrar el fenómeno al que apunta la caracterización del ser-en-el-mundo como trascendencia.

La palabra ‘trascendencia’ significa, indica Heidegger, sobrepasamiento [Überstieg]. Lo que trasciende, lo trascendente, es aquello que lleva a cabo algo así como un sobrepasar [übersteigen] y se mantiene en él. La trascendencia, entendida como sobrepasamiento, mienta un “acontecimiento” [Geschehen] constitutivo del *ser* del *Dasein*. Es decir, en cuanto perteneciente al *Dasein*, la trascendencia no es uno de los múltiples modos posibles de comportamiento respecto del ente, que pueden o no tener lugar fácticamente, sino que es “la constitución fundamental de este ente [el *Dasein*] que acontece antes de todo comportamiento”⁵. Formalmente, el fenómeno de la trascendencia como sobrepasamiento consta de tres momentos constitutivos: 1) el sobrepasamiento puede entenderse como una “relación” que se extiende “desde” algo “hacia” algo, 2) al sobrepasamiento así entendido le pertenece entonces aquello *hacia lo que* [Woraufzu] se dirige, que habitualmente recibe el nombre inadecuado de “lo trascendente”, 3) en el sobrepasamiento siempre hay *algo* que es a su vez sobrepasado.

En relación con el momento 3, Heidegger señala que “lo que es sobrepasado es precisa y únicamente *el ente mismo*, concretamente cualquier ente que pueda estar o llegar a ser desocultado para el *Dasein*, y

³ La noción aparece sólo en el importante § 69 c) de SZ, titulado “El problema tempóreo de la trascendencia del mundo”, que analizaremos en la sección 2.

⁴ Cfr. HEIDEGGER, Martin. “Vom Wesen des Grundes”. In: Wegmarken, 1929. p. 123-175. En adelante WG. El texto fue publicado por primera vez en 1929 como contribución al tomo homenaje a Edmund Husserl con ocasión de su 70 cumpleaños.

⁵ *Ibid.*, p. 137.

por tanto *también y precisamente* ese ente como el cual 'él mismo' existe"⁶. De acuerdo con ello, el Dasein, en cuanto trasciende, se sobrepasa a sí mismo como ente que existe y, a la vez, sobrepasa a todo ente que no es él mismo, pero que de algún modo se muestra o puede ser descubierto por él. En cuanto a aquello hacia lo cual el Dasein trasciende (momento 2), Heidegger señala: "Pero si el ente *no* es aquello *hacia lo que* tiene lugar el sobrepaso [*Überschritt*], ¿cómo tiene que ser determinado, buscado en general este 'hacia lo que'? Nosotros llamamos el *mundo* a aquello *hacia lo que* el Dasein como tal trasciende y determinamos ahora la trascendencia como *ser-en-el-mundo*"⁷.

El mundo es aquello hacia lo cual el Dasein, mientras existe, trasciende. En la medida en que el mundo forma parte de la estructura ser-en-el-mundo, el concepto de mundo es un *concepto trascendental*. Por ende, la comprensión adecuada de la trascendencia como ser-en-el-mundo depende de que se comprenda el mundo en un sentido trascendental, por oposición a su sentido filosófico usual. ¿Qué significa pues 'mundo' en sentido trascendental?

El término 'mundo' significa, en un sentido prefilosófico, la totalidad del ente real, donde la unidad del todo es concebida como una simple adición o sumatoria de sus partes. Frente a este modo habitual de entender el mundo, Heidegger define en *WG* cuatro rasgos fundamentales del concepto trascendental de mundo, obtenidos a partir de una serie de referencias a la historia del concepto⁸:

- 1) mundo (*kósmos*) no significa este o aquel ente, ni tampoco el todo del ente como suma de sus partes, sino más bien un *cómo del ser* [*Wie des Seins*] del ente,
- 2) este cómo determina al ente *en totalidad* [*im Ganzen*], lo cual significa que el mundo constituye, en cuanto límite y medida [*Grenze und Maß*], la posibilidad de todo cómo del ser del ente en general,
- 3) este cómo que determina al ente en totalidad es en cierto modo *previo* [*vorgängig*],

⁶ *Ibid.*, p. 138.

⁷ *Ibid.*, p. 139.

⁸ Cfr. *Ibid.*, p. 142-155.

4) este previo cómo del ser del ente en totalidad es *relativo al Dasein*, a pesar de que dicha totalidad abarca a todo ente, incluso al Dasein.

Esta caracterización preliminar del concepto trascendental de mundo tiene que ser comprendida como una reformulación del concepto ontológico-existencial de mundaneidad desarrollado en SZ. El mundo no es un ente ni la unidad de la totalidad de los entes, sino una estructura trascendental de ser, que en cada caso determina el modo (cómo) de mostrarse del ente, es decir, de hacerse comprensible (rasgo 1). En este sentido, el mundo puede ser considerado como el *horizonte significativo* (cómo del ente en totalidad), *sólo* a partir del cual todo ente puede comparecer, esto es, puede mostrarse de un modo determinado, puede mostrarse en un *cómo* de su ser (rasgo 2). El sobrepasamiento hacia el mundo como aquello hacia lo que se dirige la trascendencia implica entonces la comprensión o apertura *previa* del horizonte significativo que posibilita la manifestación del ente en totalidad (rasgo 3). En efecto, “el comprender en cada caso anticipativo-abarcante [*das je vorgeifend-umgreifende Verstehen*] de esta totalidad es el sobrepasamiento hacia el mundo”⁹.

Ahora bien, el rasgo característico esencial del concepto trascendental de mundo es el mencionado en último término: el mundo en cuanto relativo al Dasein o, más concretamente, en cuanto forma parte de la estructura ser-en-el-mundo. La siguiente cita ilustra y aclara esta determinación:

En el ser de este ente [el Dasein] *le va su poder-ser*. El Dasein es de tal modo que existe *por-mor-de sí*. Pero si en el sobrepasamiento hacia el mundo ante todo se temporiza [*sich zeitigt*]¹⁰ la mismidad, entonces el mundo se muestra como aquello por mor de lo cual el Dasein existe. El mundo tiene el carácter fundamental del por-mor-de (...) y esto en el sentido originario de que el mundo ante todo predona [*vorgibt*] la posibilidad interna de todo por-mor-de ti, de él, de eso, etc., determinado fácticamente¹¹.

⁹ *Ibid.*, p. 156.

¹⁰ Heidegger introduce en WG el verbo *sich zeitigen* sin ninguna aclaración terminológica previa acerca del significado rector de la expresión en el contexto de su concepción del tiempo como temporeidad del Dasein. Por ello, los traductores apelan al significado habitual de *zeitigen* para la versión castellana: ‘madurar’ y ‘producir’ (Cfr. HEIDEGGER, Martin. *Hitos*. H. Cortéz y A. Leyte (Trads). Madrid, 2000. n. 83, p. 135). Sin embargo, en el marco de este estudio podemos traducir con pleno sentido *sich zeitigen* como ‘temporizarse’.

¹¹ Cfr. HEIDEGGER. *Vom Wesen des Grundes*, Op. cit., p. 157.

El mundo en cuanto aquello hacia lo que el Dasein trasciende tiene primariamente el carácter de aquello *por mor de lo cual* el Dasein existe y, en este sentido, es “relativo” al Dasein. En cuanto horizonte significativo para la manifestación del ente, el mundo no es un dominio abierto que, por así decirlo, se limita a extenderse por delante del Dasein. Antes bien, en la trascendencia hacia el mundo se constituye o, mejor, se temporiza la propia *mismidad* del Dasein. Esto implica que, en cuanto horizonte de manifestación, el mundo provee al Dasein la posibilidad interna para la apropiación significativa no sólo del ente que se muestra, sino también de sí-mismo, es decir, de las propias posibilidades en las que el Dasein se proyecta. Por ello, el texto señala que el mundo predona la posibilidad interna de todo *por-mor-de*, de toda posibilidad fáctica del Dasein. “El mundo como totalidad no ‘es’ un ente, sino aquello a partir de lo cual el Dasein *se da a entender* respecto de qué ente y cómo *puede* comportarse. Que el Dasein ‘se’ da entender a partir ‘de su’ mundo, significa entonces: en este llegar-a-sí a partir del mundo se temporiza el Dasein como un *sí-mismo* (...)”¹².

La trascendencia hacia el mundo acontece entonces como la apertura previa de un horizonte significativo a partir del cual el Dasein puede comprenderse a sí mismo y a todo ente que se muestra dentro del horizonte. En este sentido, la apertura previa del mundo en cuanto tal provee el “ámbito de irrupción”¹³ para las posibilidades fácticas proyectadas en cada caso por el Dasein. Pero la trascendencia hacia el mundo, entendida como el acontecimiento que predona dicho ámbito de irrupción, no puede ser concebida a su vez como la proyección de una determinada posibilidad de existencia. Heidegger se refiere al acontecer de la trascendencia como proyección, pero no como una proyección entre otras, sino como “el proyecto originario [*ursprüngliche Entwurf*] de las posibilidades del Dasein”¹⁴. La trascendencia hacia el mundo consiste pues en el proyecto originario del horizonte o ámbito de irrupción para el pleno ejercicio de la existencia.

Por su parte el ente, en cuanto aquello que siempre es sobrepasado en la trascendencia, no es sobrepasado en el sentido de que es simplemente pasado por alto. Por el contrario, en el proyecto del mundo el ente es lo

¹² *Ibidem*.

¹³ Heidegger utiliza esta expresión para referirse al mundo: “El acontecer de la trascendencia (...) es la construcción del ámbito de irrupción [*Einbruchspielraum*] para el respectivo y fáctico *mantenerse* del Dasein fáctico en medio del ente en total” (*Ibid.*, p. 170).

¹⁴ *Ibid.*, p. 158.

que propiamente se manifiesta, es decir, lo expresamente aprehendido, lo *temático*, mientras que el mundo en cuanto horizonte es aquello que permanece siempre *atemático*. En el trato inmediato con el ente, el mundo no es aprehendido como tal, sino sólo tácitamente, permanece, por así decirlo, en el trasfondo, precisamente para que el ente respectivo pueda manifestarse a partir de él. Por ello, Heidegger caracteriza el proyecto del mundo como *sobreproyecto* [*Überwurf*]:

Pero el proyecto del mundo es, en la medida en que él no aprehende propiamente lo proyectado, también siempre *sobreproyecto* del mundo proyectado más allá del ente. Sólo el proyecto previo posibilita que el ente como tal se manifieste. Este acontecer del sobreproyecto proyectante, en el que se temporiza el ser del Dasein, es el ser-en-el-mundo¹⁵.

El fenómeno al que se refiere la caracterización del ser-en-el-mundo como trascendencia consiste entonces en el sobreproyecto originario del mundo como horizonte de manifestación del ente en totalidad¹⁶. Y en consideración de este fenómeno, Heidegger caracteriza al Dasein trascendente como *Weltbildend*, como aquel que configura el mundo¹⁷. El Dasein trascendente puede configurar o formar [*bilden*] el mundo en un triple sentido: 1) como aquel que hace acontecer el mundo, que en cierto modo lo produce, 2) como aquel que, mediante la “producción” del mundo, provee una imagen [*Bild*] originaria (el cómo), que no es aprehendida como tal, 3) como aquel que, mediante dicha imagen originaria, provee el modelo [*Vorbild*] para todo ente que se manifiesta, incluso el Dasein mismo¹⁸.

En los textos citados anteriormente, Heidegger pone en conexión, si bien de modo tangencial, el fenómeno de la trascendencia con la temporización del ser del Dasein. Esta conexión no es casual sino que tiene un fundamento sistemático. Si la trascendencia o el sobrepasamiento, concebido como el proyecto originario del mundo, constituye el acontecimiento fundamental del ser del Dasein y si

¹⁵ Cfr. *Ibidem*.

¹⁶ En WG Heidegger interpreta la trascendencia no sólo como proyecto originario del mundo, determinación a la que sin embargo le corresponde un “puesto privilegiado” (*Ibid.*, p. 165): al fenómeno pertenecen también la implicación en el ente (Cfr. *Ibid.*, p. 166-167) y el comportarse del Dasein respecto del ente (*Ibid.*, p. 168). Estas determinaciones son de hecho reformulaciones de los momentos ser-ya-en y ser-cabe de la estructura del cuidado expuesta en SZ § 41.

¹⁷ Cfr. *Ibid.*, p. 158.

¹⁸ Para un desarrollo más extenso de la idea del Dasein como *Weltbildend*, Cfr. HEIDEGGER, Martin. Grundbegriffe der Metaphysik. Welt – Endlichkeit – Einsamkeit. 2a. ed., Frankfurt a. M. 1992, §§ 64-76.

este ser se funda ontológicamente en la temporeidad, entonces la trascendencia debe tener de algún modo su origen en la temporeidad. Esto implica que la trascendencia tiene que mostrarse también como un modo fundamental de temporización de la existencia. En este sentido, “la trascendencia tiene su raíz en la *esencia* del tiempo, es decir, en su constitución extático-horizontal”¹⁹. Sin embargo, fuera de esta breve aunque importante indicación, en WG Heidegger no avanza en el examen del origen tempóreo de la trascendencia. En efecto, en nota a pie de página se lee: “la interpretación tempórea [*zeitliche*] de la trascendencia se deja completamente al margen, de modo intencionado, en la presente consideración”²⁰. Tenemos pues que abandonar WG para encontrar un análisis del origen tempóreo de la trascendencia.

2. EL CARÁCTER HORIZONTAL DE LA TEMPOREIDAD

Un tratamiento expreso y relativamente detallado de la conexión entre trascendencia y temporeidad lo encontramos en las lecciones universitarias sostenidas por Heidegger con posterioridad a la publicación de SZ²¹. Al respecto, la tesis heideggeriana es: “La trascendencia del ser-en-el-mundo se funda en su totalidad específica en la originaria unidad extático-horizontal de la temporeidad”²². “La posibilidad interna de la trascendencia es, así lo afirmo, el tiempo como temporeidad originaria”²³.

Ya en el § 69 c) de SZ encontramos una primera indicación, breve pero muy significativa, acerca del origen tempóreo de la trascendencia. En este contexto, la interpretación está guiada por la siguiente pregunta conductora: “¿cómo es ontológicamente posible algo así como un mundo en su unidad con el Dasein?”²⁴. Esta pregunta interroga por la condición de posibilidad de la apertura del mundo a partir de la aperturidad [*Erschlossenheit*] del Dasein. Por ello, lo que está en cuestión no es la determinación de las condiciones existenciales de posibilidad de *un modo* determinado de ser-en-el-mundo, sino el fundamento ontológico

¹⁹ Op. cit., p. 166.

²⁰ *Ibidem.*, nota 60.

²¹ Cfr. HEIDEGGER, Martin. *Die Grundprobleme der Phänomenologie* (1927). 3a. ed., Frankfurt a. M. 1997, §§ 20, 21. En adelante *GPh*. También *Metaphysische Anfangsgründe der Logik im Ausgang von Leibniz* (1928), Frankfurt a. M. 1978, §§ 12, 13. En adelante *MAL*.

²² HEIDEGGER. *Grundprobleme der Phänomenologie*, Op. cit., p. 429.

²³ HEIDEGGER. *Metaphysische Anfangsgründe der Logik im Ausgang von Leibniz*, Op. cit., p. 252.

²⁴ HEIDEGGER, *Sein und Zeit*. Op. cit., p. 364.

de la *unidad* de la estructura ser-en-el-mundo *en cuanto tal*, es decir, de la unidad del ser-en o aperturidad con el mundo entendido como significatividad. En relación con este problema, es preciso destacar que el concepto de significatividad no se refiere sólo a la conexión mutua de los respectos remisionales que conforman el contexto pragmático, sino que también comprende “la conexión originaria de los respectos-para [Um-zu-Bezüge] con el por-mor-de”²⁵. La pregunta por la condición de posibilidad de la apertura del mundo interroga precisamente por el *fundamento* ontológico de esta conexión originaria. El fundamento de la unidad de la significatividad, es decir, de la unidad de la estructura ser-en-el-mundo, no puede ser otro que la temporeidad, pero en un determinado sentido:

*La condición tempóreo-existencial de posibilidad del mundo reside en que la temporeidad en cuanto unidad extática tiene algo así como un horizonte. Los éxtasis no son simplemente salidas hacia [Entrückungen zu] (...). Más bien pertenece al éxtasis un ‘hacia qué’ [Wohin] de la salida. A este ‘hacia qué’ del éxtasis lo llamamos esquema horizontal*²⁶.

En este texto Heidegger introduce una nueva e importante determinación de la estructura de la temporeidad: le pertenece un *horizonte* [Horizont]. De acuerdo con ello, la temporeidad se temporiza como *unidad extático-horizontal*. La tesis heideggeriana dice entonces que la unidad extática de la temporeidad es la condición existencial de posibilidad de la apertura del mundo *en la medida en que a tal unidad extática le es constitutiva un horizonte*. A su vez, la unidad extática de la temporeidad tiene un carácter horizontal debido a que a *cada éxtasis*, en cuanto salida hacia..., le pertenece esencialmente un hacia-qué de la salida, que no es propiamente un *terminus ad quem* determinado sino un *esquema horizontal*. El § 69 c) de SZ explicita también el específico esquema horizontal que corresponde a cada uno de los tres éxtasis tempóreos:

El horizonte extático es diferente en cada uno de los tres éxtasis. El esquema en el que el Dasein viene *venideramente* hacia sí, de un modo propio o impropio, es el *por-mor-de sí*. El esquema en el que en la disposición afectiva el Dasein está abierto para sí mismo como arrojado, lo concebimos como el *ante-qué* de la condición de arrojado o, correlativamente, como el a-qué del estar entregado. Caracteriza la estructura horizontal del *haber-sido*. Existiendo por mor de sí en el estar

²⁵ *Ibid.*, p. 365.

²⁶ *Cfr. Ibidem.*

entregado a sí mismo como arrojado, el Dasein, en cuanto ser-cabe..., es también presentante. El esquema horizontal del *presente* se determina por medio del *para*²⁷.

El cuadro siguiente presenta los éxtasis tempóreos y sus correspondientes esquemas horizontales:

Éxtasis tempóreo	Esquema horizontal
Futuro	Por-mor-de sí [<i>Umwillen seiner</i>]
Haber-sido	Ante-qué [<i>Wovor</i>] de la condición de arrojado
Presentación	Para [<i>Um-zu</i>]

Cada una de las nociones que, en este estadio avanzado de SZ, Heidegger presenta como esquemas horizontales de los éxtasis tempóreos, había sido ya presentada y analizada en la primera sección de SZ. El por-mor-de sí muestra una estructura fundamental del ser del Dasein: el ser-posible en cuanto aquello que, abierto en el comprender proyectante, *le va* en cada caso al Dasein existente²⁸. El ante-qué de la condición de arrojado [*Geworfenheit*] muestra también una estructura fundamental del ser del Dasein: la facticidad, abierta en la disposición afectiva, en el respectivo estado de ánimo, y *ante* la cual el Dasein está siempre ya arrojado o a la cual está siempre ya entregado²⁹. El “para” se refiere, en cambio, a un carácter de ser del ente intramundano, más exactamente, a la estructura remisional de los útiles³⁰. En vista de este nuevo paso de la interpretación, cabe entonces preguntar: ¿en qué sentido cada una de estas estructuras ontológicas tiene que ser comprendida *como horizonte* del éxtasis respectivo? ¿Qué significa la tesis de que ellas constituyen un *esquema* horizontal? Para responder estas preguntas debemos examinar antes lo que caracteriza en general al concepto de esquema horizontal. En *MAL*, Heidegger ofrece una serie de indicaciones al respecto³¹.

²⁷ *Ibidem*.

²⁸ Cfr. *Ibid.*, p. 84, 143-145.

²⁹ Cfr. *Ibid.*, § 29.

³⁰ Cfr. *Ibid.*, § 15.

³¹ Cfr. HEIDEGGER. *Metaphysische Anfangsgründe der Logik im Ausgang von Leibniz*, Op. cit., p. 269-275.

La noción de horizonte, tal como es comprendida habitualmente, mienta el límite o la periferia del campo de visión que se tiene desde una determinada posición del “espectador”. Frente a este concepto habitual, Heidegger señala que la noción de horizonte, término que procede del griego ὀρίζειν, no está referida primariamente al ver o al intuir, sino que significa simplemente “lo que delimita, lo que rodea, la *circunscripción* [Umschluß]”³². Esta caracterización inicial orienta la interpretación heideggeriana de la horizontalidad en una dirección definida: la noción de horizonte no es determinada a partir de la intuición, de la percepción o, en general, a partir de la relación sujeto-objeto en cuanto relación de conocimiento, sino a partir del ser del Dasein como *cuidado*. El horizonte en cuanto horizonte de los éxtasis tempóreos no circunscribe el campo de la mirada “teorética” –ni tampoco el campo de la acción “práctica”–, sino el ámbito del proyecto del cuidado. En este sentido, Heidegger observa expresamente que los éxtasis tempóreos no son “ningún saber algo, ninguna conciencia, y menos aún un ver”³³.

La noción de horizonte como hacia-qué de la salida extática implica una importante determinación de la estructura de los éxtasis tempóreos. En efecto, a partir de la noción de éxtasis como salida hacia..., como “raptó” o “impulso” [Schwung]³⁴, no es posible en principio determinar aquello hacia lo que los éxtasis se dirigen o, más exactamente, aquello que los éxtasis *abren*. Sin esta determinación del hacia-qué, los diversos éxtasis no podrían ser diferenciados en su específico modo de apertura y, en última instancia, serían inconcebibles. Por ello, la concepción de la temporeidad extática sólo puede ser suficientemente comprendida cuando el respectivo hacia-qué de los éxtasis es incluido expresamente en el concepto pleno de temporeidad. Los éxtasis de la temporeidad:

(...) no son simplemente salidas hacia (...), no son en cierto modo salidas hacia la nada, sino que ellos tienen en cuanto salidas hacia (...), a causa de su respectivo carácter extático, un *horizonte perteneciente al éxtasis mismo* y prefigurado [vorgezeichnete] a partir del modo de la salida, es decir, a partir del modo del futuro, del haber-sido y del presente”³⁵.

³² *Ibid.*, p. 269.

³³ *Ibidem*. Cfr. *Ibid.*, p. 263: “Espera, retención y presentación no son sólo el modo de aprehensión del luego, entonces y ahora, el modo de la conciencia de ello, sino ellos mismos el origen; la espera no es un modo de la conciencia del tiempo, sino el tiempo mismo en un sentido originario y propio”.

³⁴ *Ibid.*, p. 265.

³⁵ HEIDEGGER. Grundprobleme der Phänomenologie, Op. cit., p. 428.

Los esquemas horizontales *pertenecen* a la estructura de los respectivos éxtasis tempóreos. Pero esto sólo implica que los esquemas están *ahí, abiertos* con la respectiva apertura extática. En este sentido, la relación de los éxtasis y sus respectivos horizontes no debe ser comprendida como un modo de la relación intencional, esto es, de la relación entre *intentio* e *intentum*³⁶. Pues, por un lado, la apertura extática del Dasein no tiene el sentido de una conciencia-de, de un *consciente* estar dirigido hacia, y, por otro lado, los horizontes no son propiamente un correlato de los éxtasis, esto es, *algo determinado* que se presenta y se opone a los éxtasis. Por ello, la trascendencia hacia el mundo, en cuanto fundada en la temporización extático-horizontal, implica también, como hemos visto, un sobrepasamiento de todo ente determinado. En y con la apertura extática, los horizontes están simplemente abiertos como aquello que *delimita* o *circunscribe* los éxtasis. “Todo éxtasis se circunscribe [*umschließt*] a sí mismo, y por cierto precisamente como ἔκστασις”³⁷. El horizonte, en cuanto aquello que circunscribe el respectivo éxtasis, constituye una suerte de “figura abarcante”³⁸, que en sí misma es indeterminada en cuanto a su contenido, que está en cierto modo “vacía”, pero que sin embargo determina de antemano el ámbito de lo que puede ser abierto o descubierto por el Dasein, de lo que en general es *comprensible* a partir de ella. Heidegger ilustra esta función posibilitante y a la vez delimitante de los esquemas horizontales del siguiente modo:

Ciertamente, en el estar a la espera mismo no se encuentra en sí nada determinado, ella [la espera] por sí no puede nunca decidir inequívocamente lo que sobre su base puede ser esperado y el modo como se lo espera. Sin embargo, la salida como tal predona [*vorgibt*] algo: lo futuro [*das Künftige*] en cuanto tal, la futuridad en general, es decir, la posibilidad pura y simple. El éxtasis no produce a partir de sí algo posible determinado, sino más bien el horizonte de posibilidad en general, dentro del cual puede ser esperado algo posible determinado³⁹.

Y así como el éxtasis del futuro en cuanto estar a la espera predona o prefigura la posibilidad en general (el por-mor-de sí) como su horizonte,

³⁶ Para la interpretación heideggeriana de la intencionalidad y la determinación de sus momentos estructurales como *intentio* e *intentum*, Cfr. Prolegomena zur Geschichte des Zeitsbegriffs (1925), 2a. ed., Frankfurt a. M. 1988, § 5, Grundprobleme der Phänomenologie, Op. cit., §§ 9, 15 y Metaphysische Anfangsgründe der Logik im Ausgang von Leibniz, Op. cit., § 9.

³⁷ HEIDEGGER. Metaphysische Anfangsgründe der Logik im Ausgang von Leibniz, Op. cit., p. 269.

³⁸ Cfr. OPILIK, Klaus. Transzendenz und Vereinzelung. München: Freiburg, 1993. p. 115: “Der Horizont, der die Ekstasen jeweilig ‘umschließt’, verleiht ihr erst eine umrissene Gestalt (Schema)”.

³⁹ Op. cit., p. 269.

también en los restantes éxtasis, haber sido y presentación, se encuentra prefigurado el respectivo esquema horizontal. En efecto, como quedó expuesto en SZ, el éxtasis del haber-sido abre en general la facticidad, la *condición de arrojado*, como su horizonte, sólo ante el cual el Dasein puede asumirse como el ente que ya es (repetición) o cerrando el cual el Dasein se olvida radicalmente a sí mismo. Del mismo modo, el éxtasis de la presentación abre de antemano la estructura de la *remisión* (el “para”) como su horizonte, sólo dentro del cual el Dasein puede descubrir el ente intramundano como útil. De esta manera, mediante el respectivo esquema horizontal queda delimitada y prefigurada la apertura de los éxtasis en una dirección específica, que distingue la apertura de cada éxtasis de la de los restantes.

En el existir fáctico, la vaciedad y generalidad de los horizontes está siempre ya determinada por lo que en cada caso queda abierto o descubierto en ellos: “Con el Da-sein fáctico queda siempre proyectado en el horizonte del futuro un poder-ser, en el horizonte del haber-sido queda abierto el ser-ya y en el horizonte del presente queda descubierto el objeto de ocupación [*Besorgtes*]”⁴⁰. Sin embargo, en cuanto figuras vacías que circunscriben los éxtasis, los esquemas se distinguen de lo que puede ser abierto o descubierto fácticamente a partir de ellos, de todo lo que puede mostrarse “a su luz”. Así, en la medida en que se distinguen de todo ente, ningún esquema horizontal es espacio-temporal, es decir, “no está localizado ni espacial ni temporalmente en sentido habitual”⁴¹. Así como los éxtasis tempóreos no son estructuras “temporales”, tampoco los esquemas horizontales son horizontes “espaciales”, sino *espacios abiertos de significación*, de manifestación significativa, de comprensibilidad para todo ente.

Ahora bien, en el espacio abierto de manifestación lo que se manifiesta es el ente respectivo, no el horizonte mismo. En este sentido, Heidegger muestra que, en el trato ocupado con los útiles, la estructura de la remisión, si bien ha sido comprendida, permanece inmediatamente atemática y es el útil respectivo o la obra lo que se muestra en un primer plano⁴². Lo mismo ocurre en relación con la autocomprensión inmediata del Dasein: si bien el Dasein se proyecta constantemente en posibilidades, el ser-posible *en cuanto tal* permanece oculto, de modo

⁴⁰ HEIDEGGER, *Sein und Zeit*. Op. cit., p. 365.

⁴¹ HEIDEGGER, *Metaphysische Anfangsgründe der Logik im Ausgang von Leibniz*, Op. cit., p. 269.

⁴² Cfr. Op. cit., § 15.

que el Dasein tiende a interpretarse pre-ontológicamente a partir del ente intramundano en cuanto ente real. Entonces, el horizonte permite “ver”, en el sentido más amplio, lo que se muestra en él, pero él mismo “no puede justamente ser, en tanto que directa y unívocamente mentado, tema de un aprehender. El horizonte tiene que ser no-temático [*untematisch*], pero no obstante ser visto. Sólo así puede convertirse en tema lo que comparece en él como tal”⁴³. El hecho de que permanezcan *inmediatamente* atemáticos, constituye pues un rasgo distintivo fundamental de los esquemas horizontales. En cuanto espacios de manifestación significativa, es inherente a los horizontes del tiempo originario permanecer inmediatamente atemáticos, a fin de que lo que se muestra en ellos pueda mostrarse temáticamente.

3. LA TEMPOREIDAD HORIZONTAL COMO ORIGEN DE LA TRASCENDENCIA

En la medida en que los esquemas horizontales forman parte de la estructura de los éxtasis tempóreos, ellos no *son*, sino que cada horizonte “se temporiza en y con el éxtasis, él es su *ekstema*”⁴⁴. En este texto Heidegger utiliza la expresión ‘ekstema’ con el propósito de poner de relieve la derivación morfológica de : κστήμα a partir de ἔκστασις, de modo análogo a como se forman σύστημα a partir de σύντασις ο σύνημα a partir de σύνθεσις. De este modo, procura subrayar también la unidad fenoménica de éxtasis y esquema. Pero, ¿qué sucede cuando la unidad fenoménica de éxtasis y esquema es considerada desde la perspectiva de la unidad de los tres éxtasis?

Futuro, haber-sido y presentación se temporizan como una unidad extática. En esta unidad extática se funda a su vez la unidad originaria de los esquemas horizontales: la temporeidad se temporiza como una unidad extático-horizontal. Y puesto que los esquemas horizontales constituyen una unidad originaria, Heidegger habla de “el horizonte de la temporeidad total”⁴⁵. Este *horizonte uno* de la temporeidad, constituido a partir de la unidad de los esquemas, es el horizonte de toda posible apertura del ser y descubrimiento del ente, esto es, “determina aquello *respecto de lo cual* [*Woraufhin*] el ente que existe fácticamente está esencialmente *abierto*”⁴⁶. En este sentido, el horizonte

⁴³ HEIDEGGER, Martin. Kant und das Problem der Metaphysik. Frankfurt a. M., 1991. p. 123.

⁴⁴ HEIDEGGER. Metaphysische Anfangsgründe der Logik im Ausgang von Leibniz, Op. cit., p. 269.

⁴⁵ HEIDEGGER, Sein und Zeit. Op. cit., p. 365.

⁴⁶ *Ibidem*.

uno de la temporeidad comprende y posibilita la co-originaria apertura del propio ser del Dasein como ser-posible en el horizonte del futuro y como ser-arrojado en el horizonte del haber-sido y la apertura del ser del respectivo ente que comparece dentro del mundo en el horizonte de la presentación. Esto implica que la unidad de los horizontes de la temporeidad constituye el *horizonte trascendental* "de toda comprensión del ser en general"⁴⁷, es decir, de toda posible apertura significativa del ser del ente por parte del Dasein.

La unidad de los esquemas horizontales así comprendida es entonces la condición *tempórea* de posibilidad de la apertura del mundo y, en cuanto tal, el fundamento de la unidad de la estructura ser-en-el-mundo, es decir, el origen ontológico de la trascendencia. "La unidad esquemática del horizonte de la temporeidad no es otra cosa que la condición tempórea de la posibilidad del mundo y de su pertenencia esencial a la trascendencia"⁴⁸. En efecto, si la unidad del ser-en-el-mundo consiste en la conexión originaria de los respetos-para con el respectivo por-mor-de sí, entonces el mundo en cuanto significatividad sólo puede ser *proyectado* en y con la apertura tempórea del ser del Dasein. Y esto se debe a que la unidad originaria de los esquemas horizontales posibilita la conexión de los respetos-para (esquema de la presentación) con el arrojado por-mor-de sí (esquemas del futuro y del haber-sido) del Dasein. La unidad de los horizontes tempóreos constituye pues originariamente el *ahí* [Da], el ámbito de sentido abierto para el pleno ejercicio del ser del Dasein como cuidado.

A partir de la aclaración de la fundación ontológica del mundo en la temporeidad en tanto que extático-horizontal, es posible también comprender en todo su alcance el sentido de la tesis heideggeriana acerca de la pertenencia del mundo a la estructura de ser del Dasein. En efecto, puesto que la apertura del mundo se funda en la temporización extático-horizontal del Dasein, no es posible afirmar que el mundo *es* en los diversos sentidos de ser de lo intramundano, es decir, el mundo no es a la mano ni tampoco es algo real⁴⁹. Antes bien, desde el punto de vista ontológico, sólo puede decirse que el mundo *existe* en el sentido de que *está ahí* [ist da], extático-horizontalmente abierto. "En la medida en que el Dasein se temporiza, *es* también un mundo (...) Él 'está ahí' [ist da] con el fuera-de-sí [Außer-sich] de los éxtasis. Si ningún Dasein existiera, tampoco estaría 'ahí' un mundo"⁵⁰.

⁴⁷ *Ibid.*, p. 1.

⁴⁸ HEIDEGGER. *Metaphysische Anfangsgründe der Logik im Ausgang von Leibniz*, Op. cit., p. 269-270.

⁴⁹ El mundo no es, sino que "mundeá" [weltet] a partir de la temporización extático-horizontal: "Lo esquemático se temporiza oscilando como un mundear [welten]" (*Ibid.*, p. 270).

⁵⁰ HEIDEGGER, *Sein und Zeit*. Op. cit., p. 365.

Heidegger se refiere al peculiar carácter ontológico del mundo también mediante la tesis de que el mundo es *nada* (*nihil*). Pero el mundo puede ser caracterizado como nada no en el sentido de un *nihil negativum*, esto es, como pura y simple negación vacía de algo. Antes bien:

El mundo es nada en el sentido de que no es ningún ente [*nichts Seiende ist*]. Ningún ente y sin embargo algo, que hay [*etwas, was es gibt*]. El 'es' [alemán], que da este no-ente, no es ello mismo ente, sino que es la temporeidad que se temporiza. Y lo que ésta temporiza como unidad extática es la unidad de su horizonte: el mundo. El mundo es la nada, que se temporiza originariamente, lo que en y con la temporización surge [*Entspringende*] pura y simplemente –nosotros lo llamamos por eso el *nihil originarium* [nada originaria]⁵¹.

Mediante esta concepción del mundo como *nihil originarium* y, en cuanto tal, como horizonte *que surge* en y con la unidad de los esquemas de los éxtasis, la interpretación de la temporeidad como origen de la trascendencia alcanza su punto más extremo. En este punto, Heidegger determina la temporeidad como *principio productivo* del ser del Dasein en un sentido ontológico-trascendental. “Así se muestra entonces la peculiar productividad interna de la temporeidad en el sentido de que el producto es precisamente una nada peculiar: el mundo”⁵². Desde el punto de vista sistemático, éste es el resultado más relevante de la introducción del carácter horizontal en la determinación de la estructura de la temporeidad⁵³.

⁵¹ Op. cit., p. 272.

⁵² *Ibidem*. El término “productividad” que Heidegger utiliza en este contexto para caracterizar a la temporeidad extático-horizontal en cuanto origen de la trascendencia procede de Kant, quien concibe a la imaginación pura como una *facultad productiva* en sentido trascendental. “Kant se ha topado por primera vez con la productividad del ‘sujeto’ en su doctrina de la imaginación productiva trascendental” (*Ibidem*).

⁵³ En la concepción heideggeriana del tiempo como temporeidad, la introducción de los esquemas horizontales tiene también otra función sistemática, que en el proyecto ontológico diseñado en *SZ* (Cfr. Prólogo, §§ 5, 6) es ciertamente la más relevante: la de posibilitar la interpretación temporaria [*temporale*] del ser en cuanto tal, es decir, la determinación originaria del sentido del ser a partir del tiempo en cuanto horizonte trascendental de toda comprensión del ser, y, sobre su base, la destrucción de la historia de la ontología. Puesto que la tercera sección de la segunda parte y toda la tercera parte de *SZ*, que debían desarrollar este conjunto de problemas, no fueron publicadas posteriormente, el mencionado proyecto ontológico de *SZ* permaneció públicamente inconcluso. Sin embargo, en los §§ 20 y 21 de *GPh* Heidegger desarrolla un análisis del esquema horizontal del éxtasis de la presentación, al que comprende como *presencia* [*Praesenz*], a partir del cual esboza una interpretación temporaria del ser como ser a la mano y de la “tesis kantiana” sobre el ser como posición. Para el análisis de la génesis textual y de la articulación temática de *GPh* en conexión con el proyecto ontológico-histórico de *SZ*, Cfr. VON HERRMANN, F.-W. Heideggers ‘Grundprobleme der Phänomenologie’. Zur ‘Zweiten Hälfte’ von Sein und Zeit. Frankfurt a. M. 1991.

Pero la productividad de la temporeidad es una productividad *interna*, en el sentido de que queda excluido de antemano cualquier tipo de “fundamento” subyacente a la temporeidad misma, que a su vez pudiera producir la predonación de los esquemas horizontales y ser el origen del horizonte-mundo. El carácter productivo y originario de la temporeidad implica que la predonación del respectivo horizonte, la delimitación esquemática de los éxtasis, tiene lugar “en y con” la temporización de la temporeidad misma. En cuanto extática, la temporeidad pre-figura, pre-delinea, sus propios esquemas horizontales. En este sentido, Heidegger caracteriza la temporización de la temporeidad como “autoproyecto originario” [*ursprüngliche Selbstentwurf*]⁵⁴. El acontecer de la trascendencia como sobreproyecto de un mundo encuentra entonces su origen ontológico en el autoproyecto originario de la temporeidad. En cuanto autoproyecto originario, la temporeidad es “*Origo der Transzendenz*”⁵⁵ o “*Ur-transzendenz*”⁵⁶, origen de la trascendencia o trascendencia originaria: el pluridimensional salir extático hacia el horizonte respectivo que *ha sobrepasado ya* todo lo que puede ser abierto o descubierto en él, todo lo que puede ser significativamente apropiado por el Dasein. “Este oscilar [*Schwingen*] de los éxtasis que se temporizan es en cuanto tal la sobreabundancia [*Überschwung*], teniendo en cuenta todo posible ente que puede ingresar fácticamente en un mundo”⁵⁷.

OBSERVACIONES FINALES

A partir del examen del carácter horizontal de la estructura de la temporeidad puede verse claramente que las fórmulas “horizonte de posibilidad” u “horizonte trascendental”, con las que Heidegger caracteriza el tiempo originario en lugares clave de *SZ*, no son usadas de modo incidental, sino antes bien elegidas por razones estrictamente sistemáticas. En efecto, en la medida en que la unidad de los esquemas horizontales delimita de antemano el pluridimensional espacio de significación abierto en la temporización del ser del Dasein, el tiempo en cuanto temporeidad extática constituye el “horizonte de posibilidad para toda comprensión del ser en general”⁵⁸. A través de esta determinación del tiempo en *SZ* quedaron sistemáticamente conectadas, por un lado, la

⁵⁴ HEIDEGGER. Grundprobleme der Phänomenologie, Op. cit., p. 436.

⁵⁵ HEIDEGGER. Metaphysische Anfangsgründe der Logik im Ausgang von Leibniz, Op. cit., p. 272.

⁵⁶ Cfr. OPILIK. Op. cit., p. 123.

⁵⁷ Op. cit., p. 270.

⁵⁸ HEIDEGGER, Sein und Zeit. Op. cit., p. 1.

interpretación ontológico-existencial de la temporeidad como origen de la trascendencia y, por otro lado, la interpretación ontológico-temporaria del sentido del ser en cuanto tal, de manera que la interpretación temporaria del ser tenía que ser preparada por una previa interpretación existencial del tiempo.

Ahora bien, la caracterización del tiempo originario como *horizonte trascendental* en SZ tiene principalmente el propósito de señalar hacia la pregunta por el sentido del ser, hacia la problemática de la temporariedad del ser [*Temporalität des Seins*], y no hacia la cuestión del origen tempóreo de la trascendencia del Dasein, aún cuando en el § 69 c) la noción de esquema horizontal es introducida de modo sumario. Por ello, el hecho de que Heidegger haya renunciado a la publicación de "Tiempo y ser" y de la anunciada segunda parte de SZ, condujo a la convicción generalizada de que la concepción de la temporeidad como horizonte trascendental de la comprensión del ser había llegado a un punto de estancamiento irreversible y estaba, por así decirlo, exhausta.

Sin embargo, como hemos intentado mostrar en el presente trabajo, lo cierto es que Heidegger no sólo no abandonó sino que, antes bien, profundizó, en sus lecciones universitarias de los años posteriores a la publicación de SZ, la concepción de la estructura extático-horizontal de la temporeidad como base explicativa para la comprensión del origen ontológico de la trascendencia y de la aperturidad del Dasein. Y a pesar del abandono progresivo de la terminología característica de la Ontología fundamental, que incluye la prescindencia de términos de inequívoca raigambre metafísica como 'ontología', 'trascendencia', 'horizonte trascendental', 'condición de posibilidad' o 'esquema', la conexión sistemática entre temporeidad y trascendencia del Dasein en el sentido aquí analizado se afianza como una posición en la que Heidegger permanece. En este sentido, dicha conexión se convierte en el hilo conductor de la reflexión heideggeriana acerca del tiempo después de la llamada *Kehre* (viraje) en el planteo de la cuestión del ser. Así, tanto en los manuscritos de 1936-1938, denominados *Beiträge zur Philosophie – Vom Ereignis*, como en la célebre conferencia de 1962, titulada *Zeit und Sein*, Heidegger piensa el tiempo en cuanto origen del *ahí* como *espacio-tiempo* [*Zeit-Raum*]. En esta fórmula, el término 'espacio' no se refiere al concepto físico-categorial de espacio, sino que mienta el ámbito abierto en y con los éxtasis tempóreos: el *ahí* o, de acuerdo con la fórmula preferida en este contexto tardío, *el claro* [*Lichtung*] *del ser*:

El espacio-tiempo como la unidad de temporización y espacialización [*Räumung*] originarios es él mismo originariamente el sitio-instantáneo [*Augenblicks-Stätte*], éste la abismal y esencial espacio-temporeidad de la apertura de la ocultación, es decir, del ahí⁵⁹.

Espacio-tiempo nombra ahora lo abierto, que se ilumina [*lichtet*] en el extenderse mutuo [*Einander-sich-reichen*] de futuro, haber-sido y presente. Solamente esto abierto y sólo ello concede su posible extensión al espacio que nos es conocido habitualmente. El iluminador extenderse mutuo de futuro, haber-sido y presente es él mismo pre-espacial; sólo por ello puede espaciar, es decir, dar espacio⁶⁰.

REFERENCIAS

- HEIDEGGER, Martin. (1925). Prolegomena zur Geschichte des Zeitsbegriffs. 2a. Ed. Frankfurt a. M. 1988.
- _____. (1927). Die Grundprobleme der Phänomenologie. 3a. Ed. Frankfurt a. M. 1997, §§ 20, 21.
- _____. (1928). Metaphysische Anfangsgründe der Logik im Ausgang von Leibniz. Frankfurt a. M.
- _____. (1929). "Vom Wesen des Grundes". In: Wegmarken. pp. 123-175.
- _____. (1969). Zur Sache des Denkens. Tübingen.
- _____. (1976). "Brief über den Humanismus". In: Wegmarken, Frankfurt a. M.
- _____. (1991). Kant und das Problem der Metaphysik. Frankfurt a. M.
- _____. (1992). Grundbegriffe der Metaphysik. Welt - Endlichkeit - Einsamkeit. 2a. Ed. Frankfurt a. M.
- _____. (1993). Sein und Zeit. 17a. ed., Tübingen
- _____. (1994). Beiträge zur Philosophie. (Vom Ereignis). 2a. Ed. Frankfurt a. M.
- _____. (2000). Hitos. Cortéz, H. y Leyte, A. (Tr.). no. 83.
- OPILIK, Klaus. (1993). Transzendenz und Vereinzelung. München: Freiburg.
- VON HERRMANN, F.-W. (1991). Heideggers 'Grundprobleme der Phänomenologie'. Zur 'Zweiten Hälfte' von Sein und Zeit. Frankfurt a. M.

⁵⁹ HEIDEGGER, Martin. Beiträge zur Philosophie. (Vom Ereignis), 2a. ed., Frankfurt a. M., 1994. p. 384.

⁶⁰ HEIDEGGER, Martin. Zur Sache des Denkens. Tübingen, 1969. p. 14-15.

ph

RESEÑAS

Crear en la Educación

Por Victoria Camps C. Editorial Península, Barcelona, 2008.

“Es claro, pues, que la educación debe ser regulada por la legislación y que concierne a la ciudad. Cuál desde ser esta educación y cómo se ha de educar son cuestiones que no deben echarse en olvido; porque actualmente se discute sobre estos temas, y no están todos de acuerdo sobre lo que deben aprender los jóvenes, tanto desde el punto de vista de la virtud como de la vida mejor, ni está claro si contiene más atender a la inteligencia o al carácter del alma”.

Aristóteles, Pol. 1337a 33

Tras un detallado análisis de las problemáticas a las que nos vemos cotidianamente enfrentados en términos de convivencia, Victoria Camps siguiendo sus más tradicionales líneas de investigación, nos presenta un libro muy actual, muy realista y, sin embargo, muy esperanzador. El libro tiene por objetivo mostrar que aún es viable creer en la educación, en su posibilidad e importancia. Un título más que dicente, sugerente, el cual destaca que, a pesar de las circunstancias cada día más negativas e irreconciliables de nuestras sociedades, es necesario que la esperanza en la educación prevalezca sobre los intereses netamente egoístas de los medios de comunicación, de la política y de la vida comercial.

El punto de partida de esta autora será un lugar común: la opinión según la cual “no creemos” en la posibilidad de educar en todo el sentido de la palabra, ya que existen miles de obstáculos que lo impiden y, a razón de ello, se considera inútil e incluso innecesario. Dicha apreciación, en muchos casos, ha truncado nuevos emprendimientos en el arte de educar. De acuerdo a Camps, el ejercicio educativo se ve afectado por distintos elementos que hacen de su práctica realmente un problema. Por ejemplo, el hecho de no estar de acuerdo sobre cuáles son los fines, los objetivos que debe plantearse la formación integral. Lo anterior se da por una simple razón: no existe debate, no hay argumentos, no hay interesados. La responsabilidad es delegada a las distintas instituciones que intervienen en la educación de cualquier niño, la familia, la escuela, los medios de comunicación, en términos generales, el medio social que le rodea; sin embargo, ninguno asume su papel realmente y, más allá de una queja, no existen propuestas y muy pocos interesados. La pregunta por la educación sigue siendo obviada por todos sus actores.

En el marco de estas reflexiones se mueve el trabajo que nos trae Camps en este libro. Cuando nos preguntamos por la relación entre ética y educación, la respuesta que ofrece Camps es la siguiente: “La educación está estrechamente vinculada a la ética entendida ésta como la formación del carácter de una persona”¹. Propone además que la educación cívica es una responsabilidad de varios agentes: familia, escuela, medio social y político. Sin embargo, entre estos la cooperación no ha funcionado del todo bien. Por ello, las acciones no han sido del todo efectivas.

Victoria Camps nos presenta un libro lleno de nostalgias por una educación que no existe y que cada vez se distancia más de lo que, a su juicio, debiera ser, pero también lleno de esperanzas. Un análisis real de la situación, comprometida con el asunto, con una propuesta educativa que se viene fortaleciendo desde sus trabajos anteriores y que hoy se ve con mayor claridad, esto es, una educación ética y política –cívica si se quiere– para ciudadanos con el fin de que estos se constituyan como seres humanos realmente comprometidos con el curso de la vida en sociedad.

La tesis que va a defender Camps a lo largo del libro podría resumirse más o menos como sigue: los procesos educativos que se llevan a cabo hoy en día tienen serias dificultades y carencias, lo cual se ve claramente reflejado en hechos como la incivilidad, la corrupción política, la apatía social, entre otros. Las causas de esta situación son muchas: sin embargo, lo primero que hay que comprender es que a pesar de lo difícil que parezca, la tarea de educar no es una tarea inútil y hay que tenerle fe (en el buen sentido de la palabra), hay que creer en ella, en su posibilidad, en las metas que se ha de proponer y en los procesos que deben desarrollar.

Evidentemente, existen problemas cuando pretendemos llevar a cabo procesos formativos en sociedades como las actuales. Diversos factores hacen de la tarea algo bastante complejo. Sin embargo, quedarse en el lamento y la queja, en la expectativa pasiva de ver cómo se nos pasa el mundo por el frente y no hacer nada a su respecto, es completamente irresponsable con el mundo mismo. Por eso, comparto con Camps la importancia de señalar esos males, comprenderlos, y procurar abordarlos con el fin de hacer que renazca la esperanza en ellos. El objetivo será promover un tipo de educación que medie entre la permisibilidad extrema que vemos hoy en día, y la rigidez que hubo en otros tiempos y que no podemos volver a permitir. Para lo cual, la

¹ CAMPS, V. Creer en la Educación. Barcelona: Península, 2008. p. 15.

autora propone tres razones que justifican la educación cívica que ella sustenta, y fundamenta un modo de afrontar las dificultades de las que ya hemos hecho mención.

1. *El ser humano es inacabado*. Nadie nace educado para vivir en sociedad. La formación es algo que se tiene que aprender en un proceso a lo largo de la vida.

2. *El ser humano debe ser educado para la libertad*. Dado que vivimos en una lógica de mercado y esto es una situación que no tiene mayor alternativa, es necesario formar ciudadanos con juicios de valor, con criterios sólidos que les permitan discernir entre intereses netamente egoístas y la vida en sociedad. El buen sentido del ciudadano debe ser formado si la pretensión es que sea una sociedad justa y la convivencia sea posible.

3. *Las leyes son insuficientes para regular la vida colectiva*. En la medida en que son abstractas no logran llenar los vacíos que se abren en medio del irremediable conflicto social. De este modo, hay que crear disposiciones en los ciudadanos que los lleven a sentirse dispuestos para la vida en sociedad.

Camps fundamenta la educación cívica para ciudadanos en estas tres razones. La educación que piensa Camps se entiende como educación moral, cívica, política igualmente. A lo largo del libro, la autora entreteje sus consideraciones a través de cada uno de los capítulos y busca, aunque no de forma evidente, responder una cuestión que trasciende sus reflexiones desde antaño. A saber, ¿para qué educar? ¿Cuál es la finalidad última que debe tener la educación para ciudadanos? Su respuesta empieza a evidenciarse hacia el final del libro: para saber ser libres, si se permite tal redundancia, para actuar autónomamente de manera responsable.

En el primer capítulo “La educación sin norte”, encontramos distintas reflexiones atravesadas por la siguiente tesis: la educación ha perdido su norte, ya que hubo un cambio de paradigma educativo que nos ha dejado sin suelo y sin criterios para promover procesos educativos. Camps introduce en este capítulo la distinción entre “educar *en* la libertad” y “educar *para* la libertad”, dos sentidos muy distintos de lo que significa educar y las implicaciones que ello tiene a nivel social. En un completo afán por distanciarnos de las antiguas prácticas represivas y dogmáticas de la educación confesional, hemos llegado al otro extremo de la balanza, la permisividad en todo proceso educativo.

El segundo capítulo se titula “Los buenos modales”. Se defiende una tesis importante según la cual la educación en general empieza por los buenos modales, elementos mínimos que coadyuvan a que exista una real disposición para aprehender ese contenido ético del que Camps ha venido hablando hace bastante, las ya mencionadas virtudes cívicas. Camps expone algunas razones por las cuales no es común educar en esta ética mínima de buenos modales: razones con las que discute y en contra de las cuales piensa argumentar. Con lo anterior, buscará mostrar la posibilidad de una educación auténtica que configure moralmente ciudadanos igualmente auténticos. De este modo, Camps llega a la definición de civismo entendido éste como “(...) la reivindicación de unas normas de conducir colectiva aceptadas por todos con el objetivo de hacer la vida en común más agradable”². La conclusión que busca defender este capítulo es que las reglas –al ser socialmente construidas– son necesarias para la vida en comunidad. Reivindicar el civismo significa, entre otras cosas, reivindicar la creencia en la posibilidad de una convivencia armónica entre ciudadanos. Igualmente remite a la convivencia en el espacio público y los elementos que deben ser de común aceptación para sobrellevar la vida en común.

Siguiendo un poco la línea del capítulo anterior, el tercer capítulo “Ir a la contra” nos muestra que existe un sentimiento generalizado de pesimismo y desconfianza hacia todo proceso educativo. Es una sensación en la cual cualquier intento por *formar* al ser humano pasa por ser inútil e ineficiente porque existen múltiples factores que le invalidan. Camps defiende la tesis según la cual aquello que verdaderamente educa es el medio social donde se encuentra el ser humano, dando por sentado que no hay teoría que valga y sobrevenga cuando la costumbre social promueve ideas contrarias. Concluye que la educación es efectivamente social y que hoy en día experimentamos un círculo vicioso en el que caso todo avance que se logra en materia educativa, la sociedad de consumo, el afán de acumular capital y la necesidad de competir lo confronta y frustra.

El tema de la felicidad es un tema obligado en la filosofía moral de Camps. Hablar de una ética de las virtudes y de la formación de un carácter moral, exige por demás abordar la cuestión sobre qué es aquello que nos hace felices tanto individual como colectivamente. Este tema es tratado por la autora en su capítulo cuarto “Que sean felices” en el cual

² *Ibid.*, p. 55.

entendemos por felicidad la voluntad de darle sentido a la propia vida, considerar una finalidad, unos objetivos, un proyecto que nos otorga ese bello placer de sentirnos bien con nosotros mismos de manera duradera. En este sentido, la educación que se propone deberá contemplar este elemento como punto de partida: deberá enseñarse a los seres humanos a vivir con algún propósito que le dé sentido a sus vidas.

En el capítulo quinto “El valor del esfuerzo” se defiende la tesis según la cual educar consiste también, entre otras cosas, en enseñar a esforzarse. Los nuevos paradigmas educativos propugnan que los procesos educativos sean más ligeros, y se defiende la libertad en los menores, cuando éstos aún no tienen una completa idea de lo que ello significa. Desde esta perspectiva se considera casi un absurdo plantear como objetivo que los estudiantes se esfuercen en conseguir lo propio. Antes bien, parten del hecho indiscutible, eso sí, de que a los niños y a los jóvenes no les apasiona por defecto el estudio, y en virtud de ello, promueven la atención en la motivación para el estudio antes que el hecho de inculcar ciertas actitudes o disposiciones en los menores para hacer las cosas. De este modo, Camps concluye que “la educación de los últimos años no ha sabido transmitir el valor del esfuerzo”³, siendo ésta la causa de los fracasos educativos que podemos ver hoy en día.

El capítulo sexto nos habla de los responsables del proceso educativo. Con el título “La educación subrogada” Camps pone de precedente que aunque la escuela es la institución por excelencia destinada para la educación de los niños y jóvenes, el medio social, la familia, los medios de comunicación, son instancias igualmente importantes para la formación de los futuros ciudadanos. La autora enfatiza en que las responsabilidades de la escuela y de la familia deben ser claramente diferenciadas ya que uno y otro lugar debe promover objetivos muy distintos entre sí.

En aras de recuperar la discusión suscitada en el capítulo primero del libro, Camps retoma en el capítulo séptimo la distinción entre “educar *en* la libertad” y “educar *para* la libertad” y reflexiona sobre el objetivo último de toda educación: “Aprender a ser libres”. Promueve la educación cívica en clave de virtudes, esto es, una educación que haga de los ciudadanos seres autónomos, con criterios de acción y decisión sólidos, con capacidad argumentativa y deliberativa. Esta educación

³ *Ibid.*, p. 96.

para la libertad contiene dos momentos significativamente importantes: enseñar a decidir y a escoger bien, e inculcar criterios para saber escoger, elementos que son explicados por la autora en este capítulo.

En el capítulo octavo la autora presenta “El valor del respeto”. Cuando hablamos de respeto, hacemos alusión a un elemento de mayor envergadura que la simple tolerancia. Es un valor que se encuentra en estrecha relación con la autoridad, aunque no es equidistante al afecto. Resulta ser un complemento para el aprendizaje a ser libre. Éste último significa “aprender a armonizar la propia libertad con la de los demás”⁴. El respeto es una categoría que permite comprender cuál debe ser el trato que merece todo ser humano, por el simple hecho de serlo, y se encuentra en estrecha relación con la dignidad de la que toda persona goza.

El noveno capítulo se titula “El ascensor social”. ¿Qué significa que la educación funcione como ascensor social? Pues, si la escuela es el lugar por excelencia donde la igualdad debe ser promovida en aras de posibilitar una ciudadanía que tenga igualdad de oportunidades, que sea una ascensor social significa que cualquier persona, sin importar su procedencia, credo, gusto, etnia, nacionalidad, debe verse promovido equitativamente respecto de los demás en la consecución y desarrollo de sus planes de vida buena. Si creemos efectivamente que la educación tiene como finalidad última la de constituir moralmente ciudadanos con disposiciones que contribuyan al interés público, hemos de creer también que la educación es la herramienta que mejor contribuye a promover los valores que son socialmente aceptados y por los cuales una sociedad justa se constituye: a saber, libertad e igualdad.

El capítulo diez se titula “Controlar las emociones” y nos habla de la importancia que tiene para la vida en sociedad, que nosotros los seres humanos controlemos nuestras emociones de acuerdo a cada circunstancia. Esta idea parte de la tesis ya aceptada por Camps, según la cual, la perspectiva de mundo que conciba una persona parte tanto de la racionalización que haga de éste, como la manera particular en cada caso, de cómo lo sentimos. La realidad nos afecta emocionalmente, por eso es importante que sepamos reaccionar ante ella. Y es porque nos afecta y nos emociona que nos vemos movidos, motivados a actuar respecto a ello con diferentes maneras de reaccionar.

⁴ *Ibid.*, p. 145.

De acuerdo con Camps, el mejor maestro para enseñar una buena costumbre es la práctica misma, el ejemplo que se dé de ella. En el capítulo once “El valor del ejemplo” se destaca que el papel de padres y maestros, y de todo contexto social que nos rodea, es significativamente importante en la formación de los menores ya que éstos observan todo lo que les rodean y aprenden de los demás, tanto las buenas como las malas costumbres. Es por tal motivo que Camps recalca la importancia que debe tener la coherencia entre lo dicho y los hechos. Para tal efecto el trabajo entre las distintas instancias educativas será completamente necesario para que la educación sea tomada *en serio*.

Vemos pues que desde una perspectiva aristotélica, Camps procura recordar la importancia de la educación cívica, con la cual sea posible constituir una ciudadanía comprometida con los intereses colectivos. Una educación para ciudadanos promovida por la ciudad misma.

KATHERINE ESPONDA CONTRERAS
UNIVERSIDAD DEL VALLE

ph

COLABORADORES

COLLABORATORS

BAS C. VAN FRAASSEN

La mayor parte del trabajo como filósofo ha sido en Filosofía de la ciencia y en Lógica filosófica, con algunas incursiones ocasionales en la Filosofía de la literatura y las conexiones entre el arte, la literatura y la ciencia. Ha enseñado en la Universidad de Yale, la Universidad de Toronto, la Universidad del Sur de California y la Universidad de Princeton. Actualmente está incorporado al Departamento de Filosofía de la Universidad de San Francisco. Es editor de la *Revista de Filosofía Lógica* y co-editor del *Journal of Symbolic Logic*. Miembro activo de diferentes asociaciones como la Asociación Americana de Filosofía y la Academia Americana de las Artes y las Ciencias, entre otras. Entre sus publicaciones se cuenta con *The Scientific Image* (Oxford, 1980), *The Empirical stance* (Yale, 2002), *Scientific Representation: Paradoxes of Perspective* (Oxford, 2008), entre un sinnúmero de artículos. E-mail: discufilo@ucaldas.edu.co

MICHAEL DAVIS

Doctor en filosofía de la Australian National University. Ha sido docente del departamento de Lenguajes y estudios socioculturales en la Universidad de los Andes, así como de la Australian National University. Entre sus líneas de investigación se encuentran: Sellars y sus seguidores contemporáneos, las suposiciones del naturalismo y del empirismo en la filosofía analítica, Rorty como ejemplar de la filosofía analítica, el pragmatismo, las relaciones entre teoría y praxis, la razón y la lógica informal, la idea del progreso, teoría política y el hombre en la sociedad. E-mail: discufilo@ucaldas.edu.co

THOMAS NAGEL

Ph.D. en Filosofía de la Universidad de Harvard (1963). Profesor de filosofía y derecho de la Universidad de New York. Se especializa

en Filosofía Política, Ética, Epistemología y Filosofía de la Mente. Es miembro de la Academia Americana de Artes y Ciencias, es miembro correspondiente de la Academia Británica, y miembro de la Sociedad Filosófica Americana. Ha recibido numerosos premios y condecoraciones y títulos honoríficos de Oxford, Harvard y la Universidad de Bucarest. Autor de *The Possibility of Altruism* (Oxford, 1970, reimpresión de Princeton, 1978), *Mortal Questions* (Cambridge, 1979), *The View From Nowhere* (Oxford, 1986), *What Does It All Mean?* (Oxford, 1987), *Equality and Partiality* (Oxford, 1991), *Other Minds* (Oxford, 1995), y *Secular Philosophy and the Religious Temperament* (Oxford 2010), entre otros. E-mail: thomas.nagel@nyu.edu

JHON RUDISILL

Ph.D. en Filosofía (2001). Se especializa en la ética social y filosofía política y teórica. Sus intereses adicionales incluyen la enseñanza de la filosofía de la raza y el género, la filosofía y la literatura y la historia de la filosofía moderna. Su investigación actual se centra en el problema de la legitimidad política y justificación en un mundo cada vez más multicultural y globalizado, y un análisis crítico de las diversas alternativas al liberalismo. Actualmente está vinculado como Profesor Asistente del Departamento de Filosofía del College of Wooster. Entre sus publicaciones están "Culture as Cure...Sometimes Worse than the Disease: A Review of Geoffrey Hartman's *The Fateful Question of Culture*" (Abril, 1999), "Review of John Christman's *Social and Political Philosophy: A Contemporary Introduction*" (Junio, 2004) y "Comment on Gerrek" (2007). E-mail: jRudisill@wooster.edu

GREGORY LAVERS

Ph.D. en Filosofía de la University of Western Ontario (2004). Su tesis de doctorado aborda la filosofía de las matemáticas de Carnap. Sus intereses incluyen la historia de la filosofía analítica, filosofía del lenguaje y la filosofía de las matemáticas. Entre sus publicaciones recientes están "*Benacerraf's dilemma and informal mathematics*" (Diciembre, 2009), "*Carnap, formalism and informal rigour*" (2008) y "*Logic and language: a philosophical perspective*" (2006, co-autor). En la actualidad se desempeña como Profesor Asistente del Departamento de Filosofía de la Universidad de Concordia en Montreal (Canadá). E-mail: glavers@alcor.concordia.ca

STEPHEN READ

Profesor de Historia y Filosofía de la Lógica del Departamento de Lógica y Metafísica de The University of St. Andrews. Su principal interés es

la noción de consecuencia lógica desde las teorías medievales de la filosofía del lenguaje, la mente y la lógica, a las preocupaciones más modernas de la lógica de la pertinencia y la filosofía de la lógica. En la actualidad se desempeña como investigador principal del proyecto de investigación 'Arché' financiado por el *Arts and Humanities Research Council* (AHRC) sobre los fundamentos de las consecuencias lógicas. Su última publicación se titula *Thomas Bradwardine, Insolubilia* (Peeters, 2010) además de un centenar de artículos en diferentes revistas internacionales. E-mail: slr@st-andrews.ac.uk

ÁNGELA ROCIO BEJARANO CHAVES

Filósofa de la Pontificia Universidad Javeriana y Licenciada en Educación Básica con énfasis en Ciencias Sociales de la Universidad Distrital. Es asistente de Investigación desde el año 2004 en el Proyecto de Investigación "Lógica Clásica y Argumentación cotidiana", avalado académicamente por la Pontificia Universidad Javeriana. Está vinculada al Grupo de Investigación De Interpretatione. Filosofía y ciencia de la interpretación, reconocido por COLCIENCIAS. Actualmente es joven Investigadora del Instituto Pensar, de Bogotá. Su última publicación se titula "El "sinónimo" de sinónimo". E-mail: annia26@hotmail.com

ELNORA GONDIM

Doctora en Filosofía de la Pontificia Universidade Federal do Piauí - UFPI (Brasil). E-mail: elnoragondim@yahoo.com.br

JORGE GREGORIO POSADA

Licenciado en Filosofía y Magister en Filosofía de la Universidad de Caldas. En esta misma universidad ha sido profesor del Departamento de filosofía orientando cursos y seminarios de filosofía moderna, filosofía de la ciencia y filosofía del lenguaje. Actualmente se desempeña como profesor de Filosofía de la ciencia y Filosofía de la Acción en el programa de Filosofía de la Universidad del Quindío. Es autor del libro *La noción tripartita del conocimiento, una introducción a la epistemología*, y de diversos artículos en revistas especializadas del país y del exterior. Sus áreas de trabajo son la filosofía de la ciencia, la epistemología y la filosofía de la acción. E-mail: gposada@uniquindio.edu.co

GABRIEL MÉNDEZ HINCAPIÉ

Candidato a Doctor en Filosofía de la Universidad Complutense de Madrid. Entre sus últimas publicaciones se cuenta con "Claves del ascenso pacífico de china y su paralelo fallido en América Latina" (2009)

y “Una revisión del concepto de explotación en el sistema - mundo de Wallerstein y su aplicación al proyecto del área de libre comercio de las Américas” (2008). Actualmente se desempeña como Profesor del Departamento de Ciencias Jurídicas de la Universidad de Caldas. E-mail: gabriel.mendez@ucaldas.edu.co

GREGORY FERNANDO PAPPAS

Ph.D. en Filosofía de la University of Texas at Austin (1990). Trabaja en la región de América Latina en sus tradiciones, la ética y la filosofía social y política. Es autor de numerosos artículos sobre la filosofía de William James y John Dewey. Su publicación más reciente es el libro *John Dewey's Ethics: Democracy as Experience* (Indiana University Press), primera interpretación integral de la ética de Dewey. Actualmente está trabajando en las conexiones filosóficas entre el pragmatismo americano y Filosofía de América Latina. Además es editor en jefe de la Revista Interamericana de Filosofía, la primera revista en línea dedicada a la filosofía interamericana con un consejo de redacción interamericano que incluye filósofos destacados de las Américas. Se desempeña en la actualidad como Profesor Asociado del Departamento de Filosofía de la Texas A&M University. E-mail: g-pappas@neo.tamu.edu

CARLOS EMILIO GARCÍA DUQUE

Doctor en Filosofía de la Universidad Internacional de la Florida. Profesor de Filosofía de la Universidad de Caldas y de Epistemología en la Universidad de Manizales. Sus áreas de interés académico son la filosofía de la ciencia y la epistemología. Autor de artículos y traducciones que han aparecido en distintas revistas nacionales e internacionales. Ha publicado los libros *Evolución histórica del pensamiento científico* (Universidad de Manizales, 1997), *Introducción a la lectura de Popper* (Universidad de Caldas, 2001) y *Popper's theory of science* (Continuum, 2006). E-mail: carlos.garcia_d@ucaldas.edu.co

PABLO ROLANDO ARANGO

Magister en Filosofía de la Universidad de Caldas. Profesor del Departamento de Filosofía de la misma universidad. Sus principales áreas de trabajo son la filosofía del lenguaje, la metafísica y la ética. Ha publicado los libros *Introducción a la filosofía moral* (Universidad de Caldas, 2005) y *De la belleza y otros caprichos del conservador* (Universidad de Caldas, 2006), así como diversos artículos y traducciones en revistas especializadas del país. E-mail: pablo.arango_g@ucaldas.edu.co

JHON FREDY LENIS CASTAÑO

Candidato a Doctor de la Universidad de Antioquia. Sus intereses están dirigidos a la Ética y la Filosofía de la Educación. Su última publicación se titula "El mal: desafío a la perfectibilidad moral y al ordenamiento jurídico. Un ensayo a partir de la filosofía kantiana" (2003). Actualmente se desempeña como docente de la Universidad de Antioquia. E-mail: johnlenisc@gmail.com

CARLOS DI SILVESTRE

Profesor (1997) y Licenciado en Filosofía (2001). Doctor en Filosofía (2009) por la Universidad Nacional de Cuyo, Argentina. Ha sido becario del Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) en Argentina y del Deutscher Akademischer Austauschdienst (DAAD) en la Albert-Ludwigs-Universität Freiburg, Alemania. Actualmente se desempeña como Profesor Asociado a cargo de las cátedras de Filosofía del Lenguaje (Seminario) e Historia de la Filosofía Contemporánea en la Universidad Nacional de Cuyo. Sus áreas prioritarias de investigación son la Fenomenología de E. Husserl, la Hermenéutica fenomenológica de M. Heidegger y la filosofía del lenguaje de L. Wittgenstein. Ha publicado reseñas, traducciones y artículos sobre temas de fenomenología y hermenéutica filosófica en volúmenes colectivos y revistas especializadas de Argentina, Chile y España. E-mail: carlosds@logos.edu.ar

KATHERINE ESPONDA CONTRERAS

Licenciada en Filosofía de la Universidad del Valle. Actualmente vinculada a la línea de investigación *Daimôn* del Grupo Praxis, Universidad del Valle. Estudiante de la Maestría en Filosofía de la misma Universidad y Joven Investigadora de COLCIENCIAS. E-mail: keclucas@gmail.com

ORIENTACIONES PARA POTENCIALES AUTORES DE *DISCUSIONES FILOSÓFICAS*

La Revista *Discusiones Filosóficas* recibe colaboraciones originales en forma de artículos que reportan resultados de investigación, reseñas, traducciones y ponencias sobre sus temas de interés (filosofía y filosofía de la literatura). Estas colaboraciones se publican en español, inglés o portugués, a menos que los autores soliciten explícitamente la traducción al español. La revista también invita a los lectores a enviar textos cortos de réplica a los artículos publicados, así como reseñas de libros especializados publicados recientemente. Las colaboraciones deben ser enviadas a su Consejo Editorial.

Los trabajos deben ceñirse a las normas establecidas para el caso (ver formato de artículo en cada edición) y estar acompañados de un resumen (abstract) no mayor de 150 palabras, preferiblemente incluyendo también su versión inglesa, y no menos de cuatro palabras clave o descriptores. Las propuestas serán enviadas en copia de papel y en disquete (programas Word). No se considerarán propuestas cuya extensión exceda las 7000 palabras (incluyendo el resumen). Todos los artículos serán sometidos a evaluación anónima por parte de lectores expertos, en el mismo orden de llegada. La versión electrónica también se puede remitir como archivo adjunto al E-mail: discufilo@ucaldas.edu.co En cada artículo se debe identificar la institución a la que pertenece su autor o una mínima información profesional e investigativa que, a su criterio, pueda ser mencionada. Los potenciales autores deben certificar que no han enviado su propuesta de publicación a ninguna otra revista y que se abstendrán de someterla a consideración de otra revista mientras dura el proceso de evaluación en la nuestra.

Las referencias bibliográficas pueden aparecer como nota de pie de página o integradas en el texto, entre paréntesis, así: (apellido del autor, año: número de página). En este último caso, deben aparecer las referencias completas al final del trabajo.

Para el fascículo correspondiente al semestre enero-junio, la revista recibe propuestas de publicación hasta el 31 de marzo. Para el fascículo correspondiente al semestre julio-diciembre, las propuestas se reciben hasta el 30 de septiembre. Los artículos recibidos después de esas fechas serán considerados para la publicación del semestre siguiente. Todos los artículos son sometidos a una evaluación doblemente anónima. Los autores serán notificados de la decisión de los árbitros en los dos meses siguientes a la recepción de sus propuestas. Sólo se publicarán los artículos que superen satisfactoriamente el proceso de evaluación.

GUIDELINES FOR POTENTIAL CONTRIBUTORS

The journal *Discusiones Filosóficas* accepts original contributions such as research papers, reviews, translations and articles on its topics of interest (philosophy on all fields and philosophy of literature). These contributions are published in Spanish, English or Portuguese, unless the author explicitly requests a Spanish Translation. The journal also invites the readers to send short texts in response to previously published articles, as well as reviews of specialized books published recently. The collaborations should be sent to the Editorial Board.

The texts should follow the established norms for each case (see article format in each edition), and be accompanied by an abstract no longer than 150 words, preferably including its Spanish translation, and at least four key words. A printed (Word program) and a CD copy of the proposals should be sent to the Editorial Board. Proposals exceeding 7000 words (including the abstract) will not be considered. All of the articles will be subjected to anonymous evaluation by expert readers, in the order of arrival. Electronic versions can also be mailed to the following e-mail address: discufilo@ucaldas.edu.co. In each article the authors must include their institutional affiliation as well as some basic professional and investigative information. Prospective authors should certify that they have not sent their proposal for publication to any other journal and that they will refrain from submitting it for consideration of another journal for as long as our evaluation process takes.

The bibliographical references can appear as foot notes or integrated in the text, in parenthesis, in the following way: (author's last name, year: page number). In the last case, the complete references should appear at the end of the text.

For the issue corresponding to the semester January-June, the journal receives potential articles until March the 31st. For the issue corresponding to the semester July-December, the journal receives potential articles until September the 30th. The contributions received after these dates will be considered for publication in the following semester. All of the articles are subjected to a double blind evaluation. The authors will be notified of the referees' decision within the two months following the reception of the proposals. Only articles that satisfactorily pass the evaluation process will be published.

w w w . 4 - 7 2 . c o m . c o



LA RED POSTAL DE COLOMBIA

› Línea de Atención al Cliente Nacional ‹
01 8000 111210

Revista Discusiones Filosóficas

FORMATO DE SUSCRIPCIÓN

Nombre / Name	
Cédula / Identification number	
Dirección / Address	
Ciudad / City	
Departamento / State	Código Postal / Zip Code
País / Country	
Teléfono / Phone Number	
Profesión / Profession	
Institución / Employer	
Correo Electrónico / E-mail	
Dirección de envío / Mailing Address	

Suscriptores Nacionales por un año. (2) Ejemplares

Se debe consignar en Bancafé, cuenta de ahorros No. 255050114 código 00HD005
Promoción e indexación de publicaciones científicas.

Mayores informes:

Vicerrectoría de Investigaciones y Postgrados
Universidad de Caldas. Calle 65 N° 26 - 10
A.A. 275 Manizales - Colombia
Tel: 8781500 ext. 11222
Fax: 8781500 ext. 11622
E-mail: revistascientificas@ucaldas.edu.co

Último ejemplar recibido / Last issue mailed:

Año/Year Volumen/Volume Número/Number Fecha / Date



Ventas, suscripciones y canjes
Vicerrectoría de Investigaciones y
Postgrados
Universidad de Caldas
Sede Central
Calle 65 No. 26 - 10
A.A. 275
Teléfonos: (+6) 8781500
ext. 11222
e-mail:
revistascientificas@ucaldas.edu.co
Manizales - Colombia



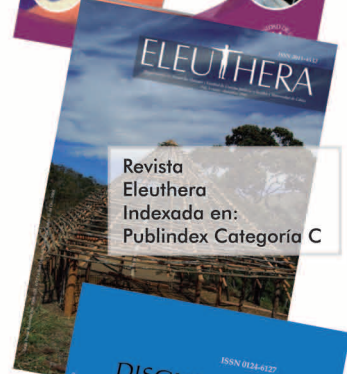
Revista
Agronomía



Revista
Biosalud
Indexada en:
Publindex Categoría B
Lilacs



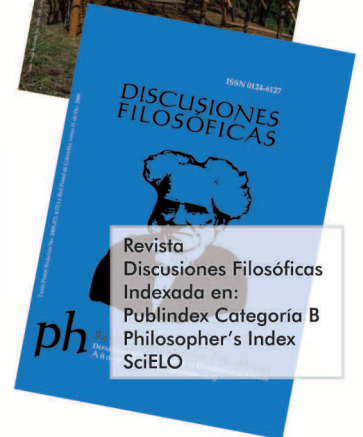
Revista
Cultura y Droga



Revista
Eleuthera
Indexada en:
Publindex Categoría C



Revista
Luna Azul (On Line)
<http://lunazul.ucaldas.edu.co>
Indexada en:
Publindex Categoría B
Index Copernicus, DOAJ



Revista
Discusiones Filosóficas
Indexada en:
Publindex Categoría B
Philosopher's Index
SciELO

Revistas





Revista Boletín Científico Museo de Historia Natural Indexada en: Publindex Categoría A2 SciELO



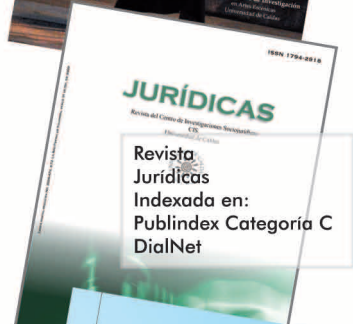
Revista Colombiana de las Artes Escénicas



Revista Veterinaria y Zootecnia



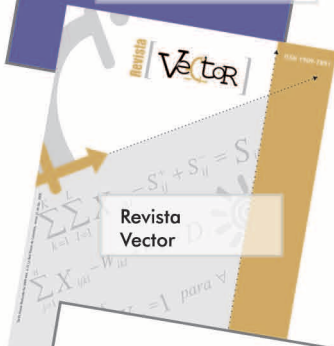
Revista Hacia la Promoción de la Salud Indexada en: Publindex Categoría B Lilacs SciELO



Revista Jurídicas Indexada en: Publindex Categoría C DialNet



Revista Latinoamericana de Estudios Educativos



Revista Vector



Revista de Antropología y Sociología (Virajes) Indexada en: Publindex Categoría C



Revista Universidad de Caldas



Revista Kepes Indexada en: Publindex Categoría C



Revista Latinoamericana de Estudios de Familia

Científicas



ph

**Esta Revista se terminó de imprimir
en diciembre de 2010 en los talleres de
Capital Graphic
Manizales - Colombia**