

REVISTA DISCUSIONES FILOSÓFICAS

ISSN 0124-6127

- Fundada en enero de 2000 -

Periodicidad semestral

Tiraje 300 ejemplares

Año 13, No 20, 298 p.

enero - junio, 2012

Manizales - Colombia

Rector

Universidad de Caldas

Ricardo Gómez Giraldo

Vicerrectora Académica

Luz Amalia Ríos Vásquez

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 reseñas 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 A2

Scientific Electronic Library On-line SciELO

Ulrich's Periodicals Directory. Latindex

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.

Universität 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

Raúl Andrés Jaramillo Echavarría - Corrector de estilo

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

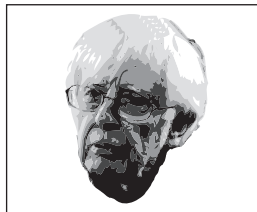


Imagen Portada:
Ilustración Rom Harré

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
SOCIAL CONSTRUCTION AND CONSCIOUSNESS <i>CONCIENCIA Y CONSTRUCCIÓN SOCIAL</i> <i>Rom Harré</i>	13
CONSCIOUS AWARENESS AND BEHAVIOUR: WHAT DISTINGUISHES CONSCIOUS EXPERIENCE FROM UNCONSCIOUS PROCESSES? <i>CONCIENCIA CONSCIENTE Y COMPORTAMIENTO: ¿QUÉ DISTINGUE A LA EXPERIENCIA CONSCIENTE DE LOS PROCESOS INCONSCIENTES?</i> <i>John Peacock</i>	37
THE POVERTY OF ECONOMISM: FREEDOM, CALCULATION, AND THE LAW <i>LA POBREZA DEL ECONOMICISMO: LIBERTAD, CÁLCULO Y LEY</i> <i>Mark A. Notturmo</i>	57
REASONS AND TWO KINDS OF FACT <i>RAZONES Y DOS TIPOS DE HECHO</i> <i>Toni Rønnow-Rasmussen</i>	95
RECONCEIVING PHILOSOPHY OF RELIGION <i>RECONCIBIENDO LA FILOSOFÍA DE LA RELIGIÓN</i> <i>Paul K. Moser</i>	115
EPISTEMIC VS ONTIC CLASSIFICATION OF QUANTUM ENTANGLED STATES? <i>¿CLASIFICACIÓN ONTICA VS CLASIFICACIÓN EPISTÉMICA DE LOS ESTADOS CUÁNTICOS ENTRELAZADOS?</i> <i>Michele Caponigro</i> <i>Enrico Giannetto</i>	137
LA IDEALIZACIÓN EN LA MATEMÁTICA <i>IDEALIZATION IN MATHEMATICS</i> <i>Thomas Mormann</i>	147

UNA DEFENSA DE LA METAINDUCCIÓN PESIMISTA <i>A DEFENSE OF PESSIMISTIC METAINDUCTION</i> <i>Daian Tatiana Flórez Quintero</i>	169
BREVES REFLEXIONES SOBRE EL DESAFÍO DEL ACTIVISTA, A LA POLÍTICA DELIBERATIVA: EL BUEN DELIBERATIVISTA, NO SIEMPRE DELIBERA <i>BRIEF THOUGHTS ON THE ACTIVIST'S CHALLENGE TO DELIBERATIVE POLITICS: THE GOOD DELIBERATIVIST DOES NOT ALWAYS DELIBERATE</i> <i>Macarena Marey</i>	187
DEL CRÁTILLO AL TRACTATUS, OBSERVACIONES A UNA TEORÍA FIGURATIVA DEL SIGNIFICADO <i>FROM CRATYLUS TO TRACTATUS, SOME REMARKS ON PICTURE THEORY OF MEANING</i> <i>Álvaro Revollo Novoa</i>	209
COMPRENSIÓN Y LENGUAJE <i>UNDERSTANDING AND LANGUAGE</i> <i>Mauricio Vélez Upegui</i>	223
FENOMENOLOGIA HERMENÉUTICA, CIRCULARIDADE E DESOCULTAMENTO <i>PHENOMENOLOGY HERMENEUTICS, CIRCULARITY AND UNCONCEALMENT</i> <i>Cezar Luís Seibt</i>	243
LA UNIÓN ALMA-CUERPO EN TOMÁS DE AQUINO EN DIÁLOGO CON A. KENNY <i>THE UNION BODY-SOUL IN THOMAS AQUINAS IN DIALOGUE WITH A. KENNY</i> <i>Patricia Moya</i>	255
DISCUSIONES	
DO WE NEED TO REFORM THE OLD T-SCHEME? <i>¿NECESITAMOS REFORMAR EL VIEJO ESQUEMA T?</i> <i>Jan Woleński</i>	273
COLABORADORES <i>COLLABORATORS</i>	286

PRESENTACIÓN EDITORIAL

EDITORIAL

Tras trece años de su fundación, la revista *Discusiones Filosóficas* llega a su número 20 con el presente fascículo. Durante los últimos años se ha consolidado como una de las revistas de filosofía más importantes del país y, gracias a su inclusión en el portal de SCIELO, su presencia en latinoamérica es cada vez mayor. En esta entrega hemos incluido catorce artículos sobre temas muy diversos que cubren la filosofía de la mente, la filosofía de la ciencia, la filosofía política, la hermenéutica y la filosofía de la religión, entre otros.

El presente fascículo se abre con el artículo "Social Construction and Consciousness" del notable filósofo Rom Harré, quien actualmente se desempeña como Director del Centro de Filosofía Natural y de las Ciencias Sociales de la London School of Economics. En este trabajo, Harré examina el carácter aparentemente variado de lo que se ha denominado en la literatura especializada el "problema de la conciencia". Harré señala que tras esta diversidad aparente se encuentra la idea de que los humanos son conscientes de que son conscientes de los objetos materiales del mundo externo, de algunos estados de sus propios cuerpos y de su capacidad de automonitorearse cuando reflexionan sobre estos asuntos. Mediante el análisis de cierto tipo relevante de conversación, Harré investiga hasta dónde se pueden construir socialmente algunos de los fenómenos comprendidos bajo el término amplio de "conciencia". En particular, el autor se interesa por su existencia previa o su existencia como resultado de tales interacciones sociales e intenta determinar si los factores sociales los pueden modificar de diversas maneras. Para realizar esta investigación, Harré parte del supuesto de que el principal instrumento a través del cual ocurren dichas creaciones y modificaciones es el lenguaje. A continuación, hemos incluido el artículo "Conscious awareness and behaviour: what distinguishes conscious experience from unconscious processes?" de John Peacock. El autor se propone analizar el concepto de conciencia consciente a través del examen de los procesos inconscientes y de los que contribuyen a la construcción de la experiencia consciente. Según Peacock, una mejor intelección de la conciencia requiere la comprensión de los criterios que separan los dos tipos de proceso. En ese orden de ideas, Peacock considera que el comportamiento nos ayuda a entender la distinción entre esos procesos y que incluso el comportamiento expresable al que corresponden los procesos inconscientes nos puede proporcionar un medio para separar la experiencia consciente.

Luego de estas discusiones sobre filosofía de la mente, hemos incluido un artículo de Mark Notturmo titulado "The Poverty of Economism: Freedom, Calculation and the Law." En este trabajo, Notturmo discute cierta información histórica sobre la relación entre Karl Popper y Friedrich von Hayek en la Viena del siglo XX, con el fin de proporcionar un contexto adecuado para enmarcar su análisis. Llama la atención el hecho de que, a pesar de tantas coincidencias sobre sus vidas, tantas afinidades políticas y concepciones similares a favor de la idea de una sociedad abierta, hasta el punto de que se les llegó a considerar aliados filosóficos, había diferencias fundamentales entre sus respectivos puntos de vista sobre la democracia, la racionalidad y el economicismo, pues los que defendía Hayek eran fundamentalmente opuestos a los de Popper. A la luz de estos planteamientos, Notturmo aprovecha las diferencias entre los puntos de vista de estos dos autores sobre el economicismo, para defender la tesis de que la crítica popperiana

al economicismo de Marx también podía aplicarse a Hayek. Tras este provocador artículo, incluimos el trabajo “Reasons and Two Kinds of Fact” del profesor de la Lund University Toni Rønnow-Rasmussen. El autor sueco plantea que las razones son hechos; es decir, que están constituidas por hechos. A partir de dicho planteamiento, que cuenta con abundante respaldo en la literatura reciente, Rønnow-Rasmussen aborda algunas cuestiones que surgen cuando aplicamos esta idea a la distinción entre razones normativas neutrales al agente y razones normativas relativas al agente. En su análisis considera dos puntos de vista sobre la naturaleza de los hechos. Según un punto de vista popular, que concibe los hechos como entidades abstractas, la dicotomía neutral-relativa no suele considerarse necesariamente problemática. Como bien se sabe, hay dos principales enfoques al respecto. El llamado *enfoque fino*, admite la posibilidad de que algunas razones *qua* hechos sean neutrales al agente y algunas sean relativas al agente, mientras el segundo enfoque, el llamado *enfoque grueso*, plantea que los enunciados de razón se refieren a entidades concretas, i.e. son entidades gruesas más que aspectos finos de tipo proposicional las que constituyen razones. El autor examina algunas de las aparentes ventajas del enfoque fino, y muestra por qué resultan problemáticas. Finalmente, explica las ventajas del enfoque grueso, que en algunos aspectos importantes resulta superior al enfoque fino, aunque aquel también enfrenta algunas objeciones importantes.

A continuación hemos incluido el artículo “Reconceiving philosophy of Religion” del profesor de la University of Loyola, Paul K. Moser. El autor afirma que los filósofos han trabajado durante mucho tiempo con concepciones de Dios que resultan inadecuadas para representar a un Dios genuinamente digno de adoración. En su opinión, la deficiencia más notoria de tal inadecuación es la omisión de la noción de severidad divina que podría resolver dicha dificultad. Moser también considera que las expectativas de los filósofos sobre Dios son equivocadas en la medida en que no concuerdan con lo que serían los propósitos relevantes del altísimo. Como consecuencia de dichas expectativas equivocadas, es común que se busquen las pruebas de la existencia de Dios en todos los lugares, menos dónde se debería adelantar tal búsqueda. Para resolver el problema, Moser sugiere reconsiderar cuidadosamente nuestras expectativas sobre Dios, y acercarnos a una epistemología de la religión, de una manera que le haga justicia a la idea de un Dios digno de adoración. El autor concluye que la evidencia disponible a los seres humanos de un Dios digno de adoración, no sería para meros espectadores, sino que buscaría desafiar la voluntad de los seres humanos para cooperar con la voluntad perfecta de Dios, como se ilustra en el caso del desafío divino de Getsemaní.

A continuación, hemos incluido un artículo de Michele Caponigro y Enrico Giannetto titulado “Epistemic Vs Ontic Classification of Quantum Entangled States?”. En este trabajo, Caponigro y Giannetto analizan, a partir de otros artículos recientes, la siguiente cuestión: ¿puede la naturaleza de los estados cuánticos entrelazados ser interpretada ontológica o epistemológicamente? Como se refleja en la literatura, los grados de libertad (y la herramienta de la partición) de los sistemas cuánticos permiten establecer una posible clasificación entre los estados entrelazados y los estados factorizables. Basados en ello, los autores italianos sugieren que la “elección” de los grados de libertad (o partición cuántica), incluso si se justifican matemáticamente, introducen un elemento epistémico, no sólo en los sistemas, sino también en su clasificación. Caponigro y Giannetto concluyen que no existen dos clases de estados cuánticos: entrelazados y factorizables sino tan sólo una clase de estados: los entrelazados, ya que los estados factorizables se vuelven entrelazados para una elección diferente de sus grados de libertad (es decir, que están entrelazados con respecto a otros observables). De la misma manera, no hay particiones de sistemas cuánticos que posean un status ontológicamente superior con respecto a otros. Por todas estas razones, ambas herramientas matemáticas utilizadas (i.e., las particiones cuánticas y los grados de libertad) son responsables de la creación de una clasificación incorrecta de los sistemas cuánticos. Finalmente, los autores sostienen que no se puede hablar de una clasificación de los sistemas cuánticos puesto que todos los estados cuánticos son estados entrelazados.

El presente fascículo continúa con el artículo “Idealization in Mathematics” del profesor de la Universidad del País Vasco Thomas Mormann. En este trabajo, el autor se propone dilucidar el papel de las idealizaciones en la evolución del conocimiento matemático, inspirado por algunas ideas de la filosofía de la ciencia y de la matemática neo-kantianas de Ernst Cassirer. Usualmente, en la filosofía de la ciencia contemporánea se da por hecho que el asunto de la idealización tiene que ver únicamente con las idealizaciones en las ciencias empíricas, en particular en la física. Pero, como bien se sabe, Cassirer sostuvo que la idealización en las matemáticas y en las ciencias tenía la misma base conceptual y epistemológica, por lo que Mormann examina su “tesis de la identidad” investigando una variedad de ejemplos de idealizaciones tomadas del álgebra, la topología, la teoría de redes y la geometría física. El autor encuentra que las idealizaciones en las matemáticas, así como en el conocimiento físico, se pueden caracterizar por la introducción de elementos ideales que conducen a compleciones. En ambas áreas, estos elementos ideales desempeñan esencialmente el mismo papel, es decir, sustituyen una variedad incompleta de objetos, mediante una variedad conceptual completa “idealizada”.

Luego hemos incluido el artículo “Una defensa de la metainducción pesimista” de Daian Tatiana Flórez Quintero, profesora de la Universidad de Caldas y de la Universidad Nacional de Colombia, sede Manizales. En este trabajo, Flórez pretende ofrecer una defensa del argumento ‘metainductivo’ propuesto por Laudan en contra del realismo epistemológico convergente. Para tal efecto, la autora analiza dos de las líneas de ataque que se han dirigido en contra del argumento metainductivo: la primera, de Diéguez, acusa a Laudan de agregar ilegítimamente un conjunto de premisas ocultas a los argumentos realistas, las cuales no son mantenidas, ni implícita, ni explícitamente por los defensores del realismo. La segunda, de Psillos, intenta reducir el número de contraejemplos ofrecidos por Laudan para mostrar que la base inductiva de su argumento no es tan amplia, ni lo suficientemente representativa como para garantizar la inducción pesimista. Flórez responde al primer ataque mediante un contraste entre los supuestos atribuidos por Laudan al realista y aquellos que Diéguez considera que Laudan atribuye de manera ilegítima al realismo epistemológico convergente. Por otro lado, Flórez muestra que una parte sustancial de la línea argumentativa de Psillos ha sido anticipada por Laudan, quien logra responder de manera categórica a sus críticas. Flórez concluye que no parece persuasivo tratar de reducir la lista y señalar que buena parte de las teorías mencionadas en el argumento no pertenecen a estadios de desarrollo maduro de la ciencia, porque -tal y como lo indica Laudan- una estrategia como ésta estaría condenada al fracaso, dado que el realista tendría que admitir que no dispone de una explicación del éxito general de la ciencia, independientemente del estadio en que se encuentren las teorías.

Tras este aporte en filosofía de la ciencia, hemos incluido un trabajo de la filósofa argentina Macarena Marey cuyo título es “Breves reflexiones sobre el desafío del activista a la política deliberativa: El buen deliberativista no siempre delibera”. En este artículo, la autora analiza las condiciones de participación en las deliberaciones que tienen lugar en política y se enfoca en el problema de cómo responder al llamado “desafío del activista” desde la perspectiva del ciudadano a quien consideramos un demócrata que delibera, aunque eventualmente pueda abstenerse de hacerlo sin perder por ello su nota distintiva. De la filosofía política pasamos a la filosofía del lenguaje a través de los aportes de Wittgenstein. Para ello hemos acogido la contribución de Álvaro Revollo Novoa, profesor de la Universidad Nacional Mayor de San Marcos en Perú quien, a partir de su trabajo “Del Cratilo al Tractatus, observaciones a una Teoría Figurativa del Significado” discute el problema relativo de cómo el lenguaje puede dar cuenta del mundo en términos representacionales desde una teoría figurativa del significado. Revollo apoya su análisis en algunos de los problemas que aparece en el diálogo platónico *El Cratilo*, considerando dos aspectos centrales: i) cómo es que, aparentemente, del conocimiento de la naturaleza del lenguaje, saltamos al conocimiento de la naturaleza de la realidad y ii) hasta qué punto el lenguaje, no es más bien un obstáculo (a la manera de un disfraz) que oculta a la realidad en vez de ser el vehículo que nos permite conocerla. Para sacar adelante sus tesis, el autor también se basa en la idea que sostiene que el lenguaje pinta o figura los objetos del

mundo. En su opinión, dicho rasgo figurativo representacional del lenguaje permite establecer el vínculo entre el diálogo platónico y el *Tractatus logico philosophicus* de Wittgenstein, sumado al aspecto que considera que existen formas (la forma de los nombres o la forma lógica), que permiten la correspondencia entre el lenguaje y el mundo, a través de sus elementos primitivos.

A continuación, hemos incluido dos trabajos del área de la filosofía hermenéutica. El primero de ellos, del profesor de la Universidad Eafit Mauricio Vélez Upegui, se ocupa de las posturas hermenéuticas de Gadamer y Ricoeur para explicar el modo como ambos pensadores parecen responder al problema que interroga por el papel que desempeña el lenguaje en la actividad humana de la comprensión, ya en el ámbito del diálogo, ya en el de la relación que media entre el acto de leer y el acto de escribir. El segundo artículo, del profesor de la Universidad Federal de Para en Brasil Cezar Luís Seibt se titula "Fenomenología Hermenéutica, Circularidade e Desocultamento". En este trabajo se muestra la forma como Heidegger desarrolla su propuesta de superar la metafísica por medio de la apertura del evento contra la objetivación que es posible por la tradición, a través de una nueva visión de la fenomenología vinculada a la hermenéutica. El autor enfatiza el asunto de las presuposiciones y las considera como precondiciones para el evento del mundo y el reconocimiento de las entidades. Finalmente, muestra cómo se puede obtener el conocimiento de manera auténtica en el círculo hermenéutico y cómo permanece escondido el horizonte del ser en el proceso de desocultamiento, aunque tanto las entidades como las precondiciones tengan lugar dentro del círculo hermenéutico.

La sección de artículos se cierra con una contribución de Patricia Moya, profesora de la Universidad de los Andes en Chile, quien nos ofrece una reflexión sobre "La unión alma-cuerpo en Tomás de Aquino en diálogo con A. Kenny". Moya examina algunos de los argumentos que el filósofo medieval expone en la primera parte de la *Suma Teológica*, concretamente el artículo 4 de la cuestión 75 y el 1° de la cuestión 76; en los cuales se ocupa de la unión del alma con el cuerpo. La autora intenta contrastar el enfoque de tipo analítico del filósofo inglés A. Kenny en los capítulos 11 y 12 de su *Aquinas on mind*, con los argumentos tomasianos basados principalmente en principios metafísicos. A partir de ello, Moya busca dar respuesta a las objeciones que Kenny plantea y a la vez esclarecer la perspectiva desde la cual Tomás de Aquino sostiene la unidad de la persona humana. En la sección de discusiones hemos acogido un artículo de Jan Woleński (University of Management and Information, Rzeszów (Polonia)), quien toma parte en la viva discusión que se ha generado desde las páginas de esta revista sobre el conocido Esquema-T de Tarski. Woleński presenta algunos comentarios críticos al análisis de Stephen Read del esquema T ofrecido en su artículo: "The Liar and the new T-scheme". El autor sostiene que Read malinterpreta los planteamientos de Tarski en algunos puntos y que el esquema A que introduce Read en su artículo, es o bien reducible al viejo esquema T (tarskiano) o bien plantea una serie de dudas esenciales.

Como ya es costumbre, quiero invitar a los lectores a seguir enviando sus sugerencias al correo de la revista y agradezco a todos los autores por su apoyo. También quiero reconocer la valiosa colaboración de David Wright, editor del Directory of Periodicals MLA Bibliography, por su orientación para que nuestra revista sea incluida en el catálogo de MLA Bibliography, a partir de la próxima actualización. *Discusiones Filosóficas* también ha sido incluida en la base de datos NewJour de Georgetown University, lo cual permitirá fortalecer su visibilidad internacional. Por último, me complace invitar a los miembros de la comunidad filosófica nacional e iberoamericana a participar en el IV Congreso Colombiano de Filosofía que tendrá lugar en Manizales, bajo el auspicio del Departamento de Filosofía de la Universidad de Caldas y la Sociedad Colombiana de Filosofía, entre el 24 y el 28 de Septiembre de 2012.

El editor

ph

DISCUSIONES FILOSÓFICAS

ph

SOCIAL CONSTRUCTION AND CONSCIOUSNESS

CONCIENCIA Y CONSTRUCCIÓN SOCIAL

ROM HARRÉ

Georgetown University, USA. rom.harre@georgetown.edu

RECIBIDO EL 25 DE FEBRERO DE 2012 Y APROBADO EL 13 DE ABRIL DE 2012

RESUMEN ABSTRACT

No parece que exista un “problema de la conciencia”, tan solo diversos enigmas que saltan a la palestra cuando reflexionamos sobre el carácter especial de los organismos superiores. Se podría decir, a grandes rasgos, que lo común a muchos enigmas sobre la conciencia es la idea de que los seres humanos son al menos conscientes de que son conscientes de las cosas materiales, de algunos estados de sus propios cuerpos y se pueden monitorear a sí mismos, pensando acerca de estos y otros asuntos.

En este artículo, utilizó el análisis de cierto tipo relevante de conversación como una forma de investigar hasta qué punto, algunos de los fenómenos comprendidos bajo el término amplio de “conciencia” se construyen socialmente. Con esto quiero decir, cuestiones como: si son traídos a la existencia o si ya existen, si son modificados de diversas maneras por factores sociales, es decir, procesos y estructuras interpersonales. Trabajaré con el principio ampliamente compartido de que el instrumento principal a través del cual dichas creaciones y modificaciones ocurren es el lenguaje.

There does not seem to be the “problem of consciousness”, only various conundrums that come to the fore when we reflect on the special character of higher organisms. Roughly one might say that common to many puzzles about consciousness is the thought that human beings, at least, are aware that they are aware of material things, some of the states of their own bodies, and can monitor themselves thinking about these and other matters.

In this paper I propose to use the analysis of certain relevant kinds of talk as a way of investigating how far some of the phenomena comprehended under the broad term ‘consciousness’ are socially constructed. By that I mean the question as to whether they are brought into being, or if already existing, are they modified in various ways, by social, that is interpersonal processes and structures. I shall work with the widely shared principle that the main instrument through which such creations and modifications come to be is language.

PALABRAS CLAVE KEY WORDS

Consciente, conciencia, gramática, lenguaje, problema de la conciencia, construcción social, programas de investigación psicológica.

Conscious, consciousness, grammar, language, problem of consciousness, social construction, psychological research programmes.

Analyzing consciousness talk

It is a useful, indeed indispensable part of the opening up of a research programme in psychology to do a little lexicography. Phenomena are made available to people, picked out from the 'hurly burly' of life (as Wittgenstein put it) by perceptual skills that are in part made what they are in the learning of a language. The rules for the use of a word can give us an insight into the domain of its application.

Thus the word 'conscious' and its cognates, though late entrants into English, are now well established, so well established that a group of people can be brought from all over the world to study whatever it is that the word 'consciousness' picks out. In doing a lexicographical exercise as a preliminary to defining a research programme we must bear in mind Wittgenstein's reminder that the diverse phenomena that are found within the domain of the words 'conscious' and 'consciousness' may not be unified by a common essence.

He thought that many ways that a common word is used is best pictured as field of family resemblances, structured by an ever shifting network of similarities and differences in the way the words at the focus of an investigation are used. While bearing this point in mind must be careful not to slip into the post-modernist exaggeration, that there is no firm ground of usage.

On the contrary symbol driven human life is possible only because among the shifting patterns of usage of words there are persisting invariants that serve generations of the users of a language to conduct their everyday lives. We must agree, as Wittgenstein remarked, in a grammar, though we must acknowledge that no grammar lasts forever. Several of the everyday uses of the word 'conscious', can be replicated by phrases making use of the words 'aware' and 'attend'. Thus the content of 'I was conscious of a disapproving atmosphere at the meeting' can be rendered by 'I was aware of a disapproving atmosphere at the meeting' while the content of 'I suddenly became conscious of a ticking sound' could be expressed with the phrase 'paid attention to'.

These groups of words are characteristically used to express a *relation* between a person and an intentional object. Their use presupposes an ontology in which the basic particulars are material things some of which are animated as persons. Examination of a larger

slice of 'aware of'/'attend to' discourse would disclose a structure in which a person served as the 'origin' of an array of entities of various sorts, patterned by the 'conscious of' relation. Such structures are typically continuously being transformed around fairly stable invariants, as one takes a walk, has lunch and so on.

But this field of family resemblances includes the abstract noun 'consciousness'. I am of the opinion that the appearance of nouns in the setting up of a research programme in psychology is a danger sign. If psychology is the science of what people *do* then verbs and adverbs ought to carry the burden of descriptive psychology rather than nouns, and especially abstract nouns. Once a word like 'consciousness', is given a central role, then we are strongly tempted to begin to look for a referent for it.

'What is consciousness?' then takes on the air of a serious question, answered by examining it closely, like a question put to a physicist, 'What is heavy water?' The fact that we cannot answer it readily makes it look as if the question is a hard one which could be answered with little more effort. But perhaps there is something wrong with the formulation of the question, drawing our thoughts off in a fruitless direction.

Let us return to our linguistic preliminaries. What do we use the word 'consciousness' for? Two uses stand out. One is simply a further candidate for replacement by one of the 'aware of'/'attend to' vocabulary. It is exemplified by such expressions 'rose to consciousness' or 'drifted into consciousness', roughly equivalent to 'became aware of' or 'noticed'. Then there are uses such as 'lost [regained] consciousness'.

Here they seem to refer to the existence of an indefinite range or domain of perception, numerous intentional objects of which someone is aware. Closely allied to this is the use of words like 'awake', said of one who is at that moment aware of the state of the material environment both inside and outside the person's body. To wake up or to regain consciousness is to come to be in such a condition as to be able to become aware of all sorts of things.

But it is not just to be aware of this and that state of affairs, but it also suggests that a certain enabling condition for being aware of intentional objects, now or once again obtains in the organism itself.

Where we were confronted with an abstract noun, used very broadly, we now have something complex but quite concrete. A person as an embodied being is in a certain state a state which is a necessary condition for coming to perceive things, which is to stand in a certain relation to them.

Prescribing consciousness research

To whom should further explorations of consciousness be entrusted? I shall argue that the study of fields of intentional objects should be the domain of discursive psychologists, while the study of the enabling conditions for the existence of such fields is the province of neuropsychology, exploring the state of the brain and nervous system characteristic of those who are awake, paying attention and so on. The first of these suggestions might look very much like a take-over bid by phenomenologists, and to a certain extent that is just what it is. The legitimacy of trying to convey to another person the nature of one's private experience by describing it needs to be defended.

The second of these suggestions might look very much like a capitulation to a reductionist programme in psychology, in which seemingly mental state terms are given a new meaning as having a neurophysiological reference. Since that is not my intention I owe the reader at least a sketch of how a phenomenological vocabulary and a neurophysiological vocabulary could be related so that the one does not displace the other.

Let us begin in the phenomenological mode. On what are perceptual domains centred? They seem to be laid out in different orientations to a person. This person is a singularity around which what is perceived is arrayed. To explore the nature of that singularity we enter into the study of selfhood, or at least of one aspect of this multivocal notion. That aspect is the sense each person has of the uniqueness of their location and trajectory through space and time, of being one and only one embodied person.

Here is a problem: do we perceive an ordered world, each from his or her own point of view because we each have an original or native sense of our own singularities? Or do we each have such a sense because we perceive the world in an ordered and *centred* way?

To encounter the structure of the perceptual domain of another person, which will include many of the same objects that one can oneself perceive, one can place one's *body* where the other person has been very shortly beforehand. Soon it is enough to imagine what the view from somewhere else would be like.

But what guarantees the authenticity of the experience as that of the perceptual domain of someone else? It can be from what they tell us about what they saw, or felt, or heard. But by what devices do we distinguish a description from the point of view of the speaker from a bland, neutral description?

The key, I argue, is in our grasp of the grammar of the first person, explicit as in 'I saw the postman go by', or implicit as in 'Can't you hear the phone?'

The expression of consciousness as awareness of something

What sort of display would be an expression characteristic of someone or some creature that was conscious? It might, for instance, be a report of how things looked from the point of view of the speaker. This is done by using the first person singular in the report. The point of view from which the world is being seen, felt, heard etc. is implicitly centred on the body of the speaker.

Using the first person in such a report also accomplishes a more complex social act, namely the taking of responsibility for what has been reported. In saying 'I can see the taxi at last' an anxious member of a travelling party not only indexes the content of the statement with his or her position in space as an embodied being, but also puts his or her character 'on the line', standing behind the report, so to say. If the plume of dust turns out to have been made by a herd of antelope and not a taxi then the reporter is liable to censure, not just for getting it wrong but for raising everyone's hopes falsely. A parrot learns to imitate the sounds of speech and croaks out 'I can see the taxi now'. There isn't one. Do we censure the parrot? Not usually. A parrot does not partake of enough of the human form of life to be liable to praise and blame for its perceptual reports.

That of which a person is conscious, that is, is aware of and/or attends to, has a structure, centred for each person, on his or her body. The fine

tuning of the perceptual centring of consciousness, in this sense, will depend on many factors, including whether the person is blind when the field of awareness is tactile and auditory, or perhaps is a patient of Oliver Sachs and suffers from some curious psycho-anaesthesia. People speak and write of the 'structure of consciousness' and most of the time they simply mean the structure of the local environment, as revealed to the visual, auditory, tactile etc. perceptual systems.

The discursive expression of that structure, as contrasted say with its expression in a landscape by the use of perspective, is usually revealed in the use of first person constructions, rather than neutral descriptions. Compare 'Look! There goes Sadie' with 'I can see Sadie!' Agentive activities, such as making a shot in billiards, could also reveal the person centred structure of the environment as perceived by the player.

Radically different uses of the first person might persuade us that the structure of consciousness was different for some person, even perhaps for some tribe if the use of an exotic grammar (from our point of view) was widespread. Where does madness end and cultural diversity begin? Only some radically different uses of first person indices would count. And just how far these uses could diverge from the customary ways of speaking is limited. In the case of people displaying multiple personality syndrome we find no evidence in their off-beat grammars of disembodied or diverse locations for perceiving the ordinary furniture of the locality.

For example there is no reason at all to think that Miss Beauchamp's or Eve White's perceptual experiences were ordered in any abnormal way. As Mary Watkins expressed the matter:

It is paradoxical that the illness (sic) of multiple personality is precisely because of its singleness of voice at any one moment, not because of its multiplicity. Improvement starts when dialogue and reflection between the selves begins to happen, when there is a multiplicity in a single moment of time, rather than multiplicity over time. (Watkins 12)

Of course the centredness of consciousness, as a field of perceived objects related to the perceiving person, cannot be multiple, even before improvement starts.

Consciousness and other kinds of awareness

So far I have discussing the grammar of 'aware of' in cases in which that awareness is perceptual. There are other uses of the phrase, more cognitive, such as understanding a meaning and drawing an inference. 'Consciousness' is so multivocal an expression that it turns up in all sorts of contexts, where it is actually doing duty for something other than the perceptual. We find feminists writing of the 'divided consciousness' of women. Examining this usage closely reveals that it does not mean anything thrilling such as 'each and every woman has two centres of awareness somehow located in the one body'. It usually means something quite mundane, that women often find themselves with two sets of (sometimes) incompatible duties; or that women often find themselves having to pay attention to more than one thing at the same time; their take home work on the family computer and the grizzling of a damp and hungry infant.

But all these objects competing for attention are located in a single material environment which the person perceives from her singular location as an embodied being. Why use expressions like 'divided consciousness'? One can only surmise that such a word is chosen for its rhetorical impact. It is a striking metaphor, but a far from a transparent terminology for a psychology of gender role differences!

Psychological research programmes

The material embodiment of human beings as persons appears in two different ways in our analysis. A place in the body, somewhere just behind the eyes, serves as the centre of the material environment including one's own body as part of it. However, I have so far not begun to develop an account of other fields of consciousness, in particular memory and imagination. The organisation of recollections and anticipations as the span of a life is patterned around a temporal origin, the 'now' of bodily existence. This too is an indexical, like the first person, and the source of the possibility of tensed statements as the public expression of thoughts about the past and the future. 'Was' means 'before now', while 'will be' means 'after now'. 'Now' is the moment contemporaneous with an act of speaking. Further research will take us deeper into the study of indexical expressions, locatives of both places and times.

Some state or pattern of states of the body seems to be a necessary requirement for all the modes of awareness that I have touched on. These are the enabling conditions as neural structures and processes. It is important that the right neural phenomena could not be identified without the use of phenomenological criteria, a point to which I will return.

From these two considerations it follows that the form of a *psychological* investigation of some aspect of human life, such as consciousness (perception) or the emotions or reasoning or attitudes or anything else, takes the following form:

1. Since every psychological process or phenomenon is made possible by virtue of a certain condition or state of the brain and nervous systems of those engaged in the activity there is a well defined though ever changing set of questions that can be put to neuroscientists. Following well established tradition I shall call the neurological foundation the 'enabling condition' for the activity. Some enabling conditions exist by nature, so to say, while others are established by training and practice.
2. Every psychological process or phenomenon is a skilled performance by a person or persons, including perception. A performance is skilled if it is intentional, that is directed to some end and is subject to criteria of correctness, propriety, in short is normatively constrained. Deeper research into the structures of these processes is facilitated by the fact that human beings have various ways of publicly expressing what is directly known privately to each individual.

These ways, since they are public, can be studied, analyzed and the principles of their construction discovered. It has been said that perception is not a skilled activity, but there are good reasons for treating it as such. For example, one must learn to identify the type or category to which an object belongs, as well as learning how to discriminate a figure from a ground. For example playing a musical instrument is a psychological phenomenon, made possible by the changes that have occurred in the brain and nervous system of someone who has acquired the skill.

The action of playing is indeed psychological, since it is intentional, the player is trying to play something, and it is normative, that is the

performance is subject to a variety of standards in reproducing the score, coordinating one's performance with those of others, phrasing 'musically' and so on. Just in the same way performing an arithmetical calculation, making a decision, displaying an emotion, forming a friendship, seeing a constellation, tasting a fine claret, and so on are all to be construed within the basic framework of enabling conditions and skilled actions. In so far as consciousness is a catch-all term for perceptions of various modes and kinds, the study of 'consciousness' too falls within this generic hybrid methodology.

Enabling conditions are somewhat more complex than the brief introductory sketch above would suggest. Not only do they include the neural state of the human organisms involved, but also there are environmental conditions that must also obtain for a skilled performance to be possible. One cannot display one's skills as a mountaineer unless there is some rock face (or artificial surrogate) to climb. One cannot be jealous unless there is someone whose rights to a good are thought by you to be inadequate relative to yours, and so on. One cannot read without some text to follow. In what follows, that is a discussion of consciousness in the discursive frame, the environmental conditions will play a minor though essential part, since it is only in a certain kind of discursive environment that consciousness as we know it, can come into existence.

The grammars of everyday life

Contemporary Anglo-American life, conceived as the joint production of meaningful and normatively constrained patterns of action, of which verbal intercourse is one species, seems to be shaped by four main 'grammars', that is more or less coherent systems of norms and semantic rules.

There is the *Soul or S-grammar*, which makes use of concepts like 'God', 'soul', 'sin', 'redemption' and so on. This grammar, once universal in the 'West', is now confined to certain rather restricted groups, tribes and regions where it was once dominant. It is still part of working practices of Mormon Utah, for instance. A main feature of S-grammars is the independence of the prime source of activity, the soul, from any necessary relation to the body it 'inhabits'.

There is the *Person or P-grammar*, in which persons are the basic particulars and originating sources of activity. It comprises the tribal dialects and idiolects of everyday life in all those regions of the world where some measure of individualism prevails and where the peculiar blend of Roman and Anglo-Saxon legal traditions are in place.

Among some of the specialised dialects of this generic grammar are the idioms of the courtroom, of the shrill claims for rights of the 'entitlement society, and so on. A main feature of P-grammars is the conceptual tie between singular embodiment and personal uniqueness and individuality.

There is the *Organism or O-grammar*: Current Western discourses make use of a third grammar, that in which the basic powerful particulars or active beings are organisms. While it has, so to say, its natural domain of application in discussions about animals it has some important uses in discourse about human beings.

The use of this grammar is becoming very widespread in reports of and discussions of the role of a human being's genetic endowment in the pattern of one's life. In so far as one's actions are the product of neural mechanisms of genetic origin one's life is embedded in a discourse in which the concept of 'person' is replaced by that 'organism' in discussions of human beings.

Animals are agentic and act teleologically, but the person grammar is extended to them only with difficulty and usually only metaphorically. Animals do not act intentionally in the full sense that would bring into play the grammar of responsibility attributions except in rare cases. Responsibility talk addressed to family pets is surely metaphorical.

When addressed to certain primates, such as domesticated chimpanzees it may have a deeper significance, widening the scope of the domain of moral agents. Though babies act for an end they surely do not act for a purpose. If Vygotsky is right the use a P-grammar with its load of 'responsibility' concepts, for talking about and usually to, neonates, is part of the necessary conditions for a human organism to mature into a person.

There is the *Molecular or M-grammar*, in which molecules and molecular clusters are the basic particulars and originating sources of activity. Among the dialects shaped by M-grammar is human physiology and

molecular biology. Discourse framed in this grammar includes such attributions of agency to molecules as the power (alleged) of melatonin to put one to sleep in the sense of a change in the brain rhythms, and excess stomach acid to cause heartburn, in the sense of discharges in the pain receptors.

Unlike the P-grammar the M-grammar is strongly hierarchical and displays emergent properties at every level. Why did top tennis players eat a banana between sets? This query requires an answer couched in M-grammar. But why they no longer seem to do so may need the P-grammar. And so too does an account of the use of cortisone to reduce the inflammation in a cartilage, and so on.

Thus we have a loose cluster of grammars that set the standards of proper discourse for the human domain, the S-, the P-, the O- and the M-grammars. Each has variants, and in certain circumstances they fit together into hierarchies, and, in other circumstances, they complement one another.

Not only does each of these grammars have its associated ontology, or catalogue of sources of activity, but each also comprises taxonomies of dependent particulars, such as action-types in the P-grammar.

These give us criteria for partitioning the flow of the activities of persons into meaningful acts. In the O-grammar we have a classificatory resource for identifying behaviour-types as partitions of the activities of pets and some of the wild animals well known to us.

The M-grammar allows for the identification of chemical reactions in organ systems, for example the interaction between carotene and cholesterol. Organs, in the M-grammar, are treated as functional partitions of the hierarchical clustering of molecules.

Each grammar has its own distinctive principles of sequence and order among basic and dependent particulars. In P-grammar these include semantic and syntactic rules, moral imperatives and story-lines. Thanks to the work of the ethologists we now see the lives of animals teleologically.

This is expressed in the O-grammar in terms of repertoires of actions directed towards maintaining their forms of life, leading to a complex repertoire of means/end principles. But in M-grammar there are only causal laws.

S- and P-grammars differ from O- and M-grammars in the way that responsibility is dealt with. This is particularly important for a philosophy of psychology, since the transition from infancy to maturity of a being that has native agentive powers and acts teleologically, occurs along the dimension of growing responsibility for what it does.

Shaver (1985) has proposed an analysis of responsibility dimensions that will do very well as a working grammar for much of the P-grammar of current English language folk psychology¹. The attribution of responsibility according to Shaver runs as follows:

A judgement made about the moral accountability of a person of normal capacities, which judgement usually but not always involves a causal connection between the person being judged and some morally disapproved action or event. (Shaver 66)

The 'causal connection' presupposes the agentive powers of a person, not a Humean regularity of stimulus-type and response-type. Hart (1963), analyzing that variant of the P-grammar that is to be found in English law, cites three necessary conditions for attributions of responsibility:

- i) That the person understand what is required.
- ii) That the person has deliberated on the matter in hand.
- iii) That the person conforms to the result of the deliberation.

One can see that in the S-grammar the concept of 'sin' depends on very similar conditions obtaining.

Another striking way in which the S- and P-grammars differ from the O- and M-grammars, is in everyday discourse of remembering. Only people (or souls) remember neither their brains nor the molecular structures that ground our recollective capacities.

Why is this? To say 'I remember' ... is to claim some kind of authority, to commit myself to what I assert about the past. How my claims to remember are taken depends on my moral standing in the community, my reliability. It engages my personal qualities. It does not refer to me as an organism.

¹ I owe notice of Shaver's work to M. A. Spackman.

Playing tennis is another example that requires the P-grammar. The exchange of shots is constrained by conventions of meaning: 'On the line is out'; and of procedure: 'Change ends after four games'. Scores accrue to people and it is people who play shots, good and bad, for which they are responsible, neither their bodies nor their racquets.

Finally I draw attention to the grammatical fact that the experiences of falling asleep and of waking, of being in pain and so on, are described in the P-grammar, since it is *persons* who notice what is going on, and its persons who suffer.

How is a phenomenology of experience possible?

The upshot of the discussion so far is that the 'study of consciousness' is a hybrid project, requiring intensive work in the phenomenology of perception, proprioception in all the sense modalities together with a study of acts of recollection, *and* requiring intensive work on the neurophysiology of the conditions which enable acts of perception and recollection.

The former are done by people, the latter happen in brains. But if the phenomenological project is to be pursued an age old problem must be solved: how is it possible for there to be a common, public language by the use of which people can discuss how the world appears, how their bodies are, and so on, each from his or her own point of view?

The P-grammar of descriptions and avowals

Many experiences are not shared with others. While we both see the same tennis ball only I feel the pain of my 'tennis elbow'. While we both see the same slice of Tin Roof Fudge Pie only I enjoy the exquisite sensations since you foolishly chose carrot cake. Yet I can discuss the pain of my tennis elbow with my doctor, and I can mortify you with my panegyric on the gustatory delights of a famous 'desert'.

But there seems to be an insoluble impasse in philosophy in which the authority of experience clashes with the conditions for sharing a meaning with someone else. In the course of criticising the idea that a language could be established on the basis only of private and personal referents for the teaching the meaning of words, Wittgenstein also provided the means for resolving the impasse.

Wittgenstein's Private Language Argument (PLA)

The shape of the argument

The 'argument' involves three interwoven and mutually supporting strands that are relevant to the distinction between descriptions of states of affairs and avowals of experience. One strand is a demonstration that the general principle that learning the meaning of a word is achieved by a teacher pointing to an exemplar cannot account for how words for private experiences can be learned.

There is no public exemplar to which the teacher can point and learner can attend. Meaningful words for sensations and bodily feelings could not have been learned by 'ostension'. Furthermore our experiences, both private and public, are ordered around spatial and temporal 'I-poles', to use Husserl's useful phrase, a feature of experience that contributes massively to the human sense of self.

Again there is no public exemplar for learning the words, such as first person pronouns, that we use to express the structured or centred 'shape' of our personal experience. Wittgenstein's demonstration that there must be some other way to learn these important classes of words than by pointing to exemplars depends upon a commonsense but fundamental observation. The words for discussing publicly the nature and structure of private and personal experience are learned with ease and used without a qualm.

The assumption we have just criticised, about the role of exemplars in learning words, would make no sense in the case of private experiences, without the further assumption that all meaningful words denote *objects*. The second strand of the PLA is an argument to show that feelings are not objects in the sense that the objects that are pointed to in ordinary situations of ostensive learning, say 'potato', are objects.

This argument is based on a reflection on the criteria by which judgements of 'same feeling' are made in intrapersonal and in interpersonal comparisons. According to Wittgenstein the concept of 'sameness' for feelings; cannot be analyzed along the same lines as we analyze concepts of sameness for such *things* as potatoes.

Judgements of 'sameness' in material contexts are based on criteria of qualitative and numerical identity. Two things are qualitatively identical

if they have more or less similar properties. Something is numerically identical, 'still the same' in different circumstances, if its integrity and continuity of existence is uninterrupted.

In some of later paragraphs in the exposition of the PLA Wittgenstein shows in some detail that these criteria of identity have no place in language games in which we make such judgements as 'I feel just as bad as I did yesterday' or 'I know just what you are feeling. I felt the same when I had finished my exams'.

Yet everyone makes lots of everyday unproblematic judgements of sameness and difference of their own and other people's feelings. Since they do not meet the criteria for 'sameness' and 'difference' of material things, bodily feelings are not objects in the relevant sense. But even if they were their privacy precludes their use as exemplars for denotational learning.

How to resolve this seemingly intractable bouquet of problems? We use feeling words perfectly well, yet it seems that we are never in a position to learn them. We talk about feelings yet they are not at all like the entities of the material world about which we also talk.

The impasse comes about because we took two things for granted. We assumed that words are all learned by the teaching pointing to an exemplar of the kind of object talked about while the pupil attends to it. We also assumed that feelings were a special sort of object or entity, for which the usual, thing-related criteria of identity, perhaps adjusted a little, were appropriate. But feelings are not thing-like. So we have two gaps to fill in understanding how it is possible to discuss our private feelings with others and to reflect on what we felt in the past. How is it that words for feelings do get their proper meanings? And what do we mean by 'same feeling' in both intra and interpersonal contexts?

In a quite similar way an intractable problem arises about that to which the word 'I' refers when I use it of myself in reporting how things are for me. How could such an expression ever get established in public language if it could be learned only from a public exemplar?

There is no public exemplar of the centredness of experience, nor is there any *thing* to which attention could be drawn. My 'self' is no more a thing-like entity than is my pain. And again we can see this by comparing my sense of my own personal identity with my sense of yours.

I think you are the same person I saw yesterday mostly by noticing the similarity of certain salient features of your appearance today as yesterday. Of course I might be mistaken. I do not have to use any criterion to judge if I am still the same person I was yesterday, and certainly not one based on bodily appearances. It does not make sense to declare: 'All along I was wrong about the person I am. I have discovered I am you'.

There are subtly different cases of uniqueness of point of view here. I can't have your pains in the way I can't sneeze your sneezes, to borrow another of Wittgenstein's images. To be that instance of pain it has to be your feeling. Thus only you can have it. I can see a herd of cows from your point of view by simply standing where you were a moment ago, but the second term of the relational structure of the seeing of the material environment is uniquely you or uniquely me.

But in both cases we slip into seemingly intractable puzzles by taking for granted that it is objects that are in question, where 'the feeling' or 'the self' are assumed to obey the same grammars as 'the colour' and 'the driver'.

Wittgenstein's suggestion that solves the linguistic problem

The suggestion that allows us to transcend the impasse is to be found in the following remarks:

How do words *refer* to sensations? -There does not seem to be any problem here; don't we talk about sensations every day, and give them names? This question is the same as: how does a human being learn the meaning of the names for sensations?-, of the word 'pain' for example.

Here is one possibility: words are connected with the primitive, the natural, expressions of the sensation and used in their place. A child has hurt himself and he cries; and then adults talk to him and teach him exclamations and, later, sentences. They teach the children new pain behaviour.

'So you are saying that the word "pain" really means crying?' -On the contrary: the verbal expression of pain replaces crying and does not describe it. (Wittgenstein 244)

There are plenty of language games which have the right combination of natural expressions and alternative verbal expressions. A child picks up a toy and chortles.

A footballer writhes in the aftermath of a vigorous tackle. Gaza weeps with disappointment and frustration when he misses the vital shoot out against Argentina. These are natural expressions of how one is feeling.

Words for feelings are learned as alternatives to natural ways of expressing feelings. Instead of chortling we learn to say 'I'm delighted'. Instead of weeping some footballers will tell you that they are 'as sick as a parrot'.

So here we have the essential move in a resolution of the problem of how it is possible to learn how to use a word for a private feeling in a public context. We learn to use a verbal formula as a substitute for a natural expression. Just the same holds for the centredness of 'consciousness' and the use of the first person. In the same language games we can find the groundings of judgements of 'same and different feeling'.

They are not made by a comparison between your feelings and mine, or between my feeling of yesterday and my feeling of today. They are made by attention to parallel patterns in the language games in which our feelings are publicly expressed.

Wittgenstein goes on to develop a thoroughgoing distinction between descriptions of public objects and states of affairs, and avowals of 'how it is with me', the form and content of private experiences. In the former case there is room for evidence, and error. In the latter there is no epistemological gap between the groan and the feeling, nor between the centring of my personal environment on my body and the use I make of 'I' in reporting on that environment from my point of view.

The groan expresses the feeling. It does not describe it. Feeling and the disposition or tendency to groan, sigh, rub the spot, weep and so on is integral parts of the same phenomenon.

Abstract any of them and the phenomenon disappears. If we have no tendency to groan then whatever the feeling is, it cannot be pain. Expressive ways of using language depend on a pre-existing ethological repertoire of natural expressions of the ways we are feeling, sad, in

pain, happy, and so on. This is part of the natural history of humankind (Robinson and Harré).

The message of this paper is that the same treatment should be given to reflexive expressions like the first person pronouns and verb inflexions as Wittgenstein gave to words for feelings. They are expressions of the centred layout of what we can perceive, of what we can recollect and so on rather than devices used to refer to a mysterious entity within the person. In the Cartesian account of personal identity the inner 'ego' is not only the centre of experience but also a substance of which our experiences are properties.

The word 'experience' springs easily to the pen in discussions like this, but it too is fraught with the possibility of bad grammatical analogies, as if the experience of something were a different phenomenon than the thing which we see, feel and so on.

We owe to Wittgenstein an all important insight for a methodology for psychological research, that the ontological distinction between thought (ineffable and meaningful) and language (audible, visible and tangible and meaningful) does not matter in certain key cases, because in the case of avowals, there is a holistic unity. Part of the art of psychological method is to be right in distinguishing those psychological phenomena for which the holistic principle holds and those for which it does not. There is no epistemological gap between a feeling and the expressions of a feeling, in those cases in which there is a natural expression, a part of human ethology, from which the substitution of words for non-verbal modes of expression can take its start.

In the case of perception and proprioception the natural expression is to be seen in the way actions are taken -- returning a serve, picking up a spoon, scratching an itch and so on. The P-grammar includes the rules for the use of first person indexicals and these are the basis of verbal substitutes for the natural expression of the structure of the environment as we perceive it.

The grammar of the expression of the environment as I perceive it

What is the role of the first person singular in everyday talk? At first glance it might look as if 'I' was a queer kind of name, one that each

person could use to refer to him or herself. But already we can see that there is something odd about that suggestion. Unless the person addressed is aware of who is speaking the self-referential function of 'I' is ineffective.

So who is 'I' for the moment is discovered, so to say, by knowing who is speaking (or writing). In this way it is quite unlike a proper name.

I know who is being referred to when someone uses the words 'John Lennon', even though he is long since dead and gone and only his songs live on. I am addressed by my proper name, but never by 'I'.

The pronoun of address is 'you', but that is not a name either. Neither pronoun is used to refer to anybody in the way a proper name can be used to refer to someone.

One common use for the first person singular, in English, is to take responsibility. 'I'll look after the children' is so close to a promise that it would usually be taken to be just another way of promising.

However, in order to know where the responsibility lays the interlocutors must know who is speaking. Immediate presence is not necessary. One can make a commitment that is activated only in the future (a will) and one can make promises over the phone. Embodiment is not a necessary condition for this use of 'I'. Commitments could be entered into at a séance.

However there is another use of 'I', often going along with the one just described. 'I' indexes the perceptual content of a report with the spatial position of the speaker.

This is common in the everyday life of positioning oneself in space. It is also the core of autobiography. Telling about an event in the first person places the speaker at or near the event in question. 'I watched the planes circle and come in to drop their bombs on the bridge' carries at least the conversational implicature that 'I was there!'

The archetypal autobiographical anecdote has both place and time indexes: 'When I was in Nevada ...' Authority for myself as the author of the story is claimed just by the use of the first person. Compare the fairy tale opening 'Once upon a time there lived a King who had three sons' with the autobiographical 'I was one of three brothers'.

Places are locations in physical space. They are defined by relations between material things. In so far as I have a place in space I have it by virtue of the fact of my embodiment. In so far as the tenses of the verbs in my utterances relate what I am saying to the times of events as past or future, my utterances themselves must be events in the same time frame as those events. In so far as those events are happenings in the material world, so must my utterances be. At least one aspect of myself must be a trajectory of connected locations in space and time.

But there is more. To be a person is to be a singularity, to have just one trajectory in space and time, to be embodied in just one material organism. This is not a ubiquitous fact that students of human life have discovered. It is part of the grammar of the 'person' concept.

Strawson (1964) pointed out that it is only as embodied beings that people are routinely identified and reidentified as the singularities they are. They share this metaphysical property with the material things among which they are embodied.

As a grammatical remark about people it fits well with the observations I have made above about the indexical force of the first person singular in reports how the environment strikes me and autobiographical narratives. Personhood is so bounded by the singularity of each human being's embodiment that neither claim to more nor to less than one person per body is permitted to stand.

The former exploits the distinction between the use of 'I' to take responsibility and its use to index reports with their spatial and temporal location. If there is a second 'I' in use by this body, then it may be taking responsibility for actions performed by the first 'I' who is then no longer to be blamed for them.

The latter is so transparently a ploy if taken contemporaneously that it seems to turn, not on the question of material location as the centre point of perception, but on issues of responsibility. Fascinating though the issue is it is not central to my analysis in here.

In summary then, in order to understand a perceptual report I have to know where the speaker is in space. But that spatial position is fixed by the relations that the speaker bears to the things in the material environment. To have a place in such an array a speaker must be embodied, as a thing amongst things.

The acts performed by the active person with whom we are engaged are embedded within sequences of events in the material environment, such as the risings and settings of the sun, the ticking of clocks and so on.

So a speaker's acts must be carried by events in the material world. To understand a perceptual report fully I must know how the event of making the report and the event reported are related in time. Both must be material events, and this again requires that the speaker be embodied.

The general principle that acts must be sustained materially has long been accorded a place in social psychology and ethology in the act-action-behaviour distinction. Behaviours are material phenomena, actions are material phenomena seen as intended by the actor, and acts are the social meanings of actions as they are seen with unfolding social episodes.

The traditional 'mind/body' distinction is no longer required to make sense of the distinction between private and public activities. Instead we are using a distinction between personal and impersonal points of view on the immediate and the distant environment. Singular embodiment, one person per body and one body per person, focuses our attention on that which is localised in space and time. People are many and history is rich in events.

The emphasis I have placed in embodiment ought not to obscure the importance of studying cognitive phenomena that are generated in the symbolic interactions of crowds and in the historical unfolding of long term 'conversations'.

If consciousness talk is nothing other than a way of talking about what one perceives from one's own unique point of view, then the 'structure of consciousness' is nothing other than the structure of the field of perceived entities from the point of view of the perceiver.

There is a singular but imperceptible 'origin' and a 'pencil' of relations to an ordered array of objects, both within and without the body. A public expression of this private and personal structure is to be found in the indexicals of the P-grammar with which we publicly make available to others 'the environment for me'.

Is consciousness socially constructed?

This sounds like a thrilling question. However, looking more closely at what might be involved it loses its exciting character. The mark of a socially constructed aspect of mind is variation in different cultures, in particular in the different ways languages perform the same or similar functions. Given that everyone's perceptual system, say as described by Gibson (1966), is inherited with certain built-in capacities in what ways could our uses of that system differ?

Fine structure

In addition to the indexical pronouns the grammar of demonstratives also reveals ways that the 'environment-for-me' is structured. English makes do with 'this' and 'that', for things seen, heard and so on near and far from the speaker.

The language of the Maya of Yucatan is richer in demonstratives. For instance 'that-seeable-by-you-but-not-by-me' is lexically distinguished from 'that-seeable-by-both-of-us'. In short Maya demonstratives facilitate joint attention to the finer structure of an array of seen objects than does English. In this mundane sense we could say that the difference between Maya consciousness and English consciousness is socially constructed, in that the language game of referring to distant objects is learned in social practices, and what is learned during the acquisition of the English demonstrates is different from what is learned when acquiring Mayan.

Salience

Even less thrilling is the commonplace that our capacities to differentiate objects in our environment is fine tuned by learning whatever discriminatory criteria are necessary to manage successfully in some local form of life. For instance the ethnobotany of one tribe will differ from that another, so what members of each will pay attention to in their environments will differ.

Further steps in the analysis

There is at least one major issue raised by the treatment offered in this paper which must remain 'unfinished business'. Eliminating 'consciousness' talk in favour of more mundane ways of expressing

and commenting upon what people can perceive both inside and outside the envelopes of their bodies, and interpreting the first person as indexing avowals with the location of the speaker, presumes that the *distinction* 'speaker-non-speaker' is in place.

This might be the 'I/thou' distinction of some theologians, or it might be the Kantian double synthesis of 'self-as-perceiver' over against 'object-as-perceived'.

Is the maturing infant built to make this distinction or is it created in Vygotskian psychological symbiosis with beings who have already acquired it? Infants and chimpanzees, though at first unable to realize that a spot of paint on their foreheads, visible in a mirror, is on their own faces, soon do realize this, and try to pick it off. Is the transition due solely to maturation of the perceptual system or is it, in some measure, socially constructed? I believe the answer is yet to be found.

Since the P- and O-grammars are both required to delineate a research programme in the psychology of perception, the question of their relationship needs to be tackled. I have taken for granted that the criteria by which organs and organ systems relevant to perception require distinctions drawn in the P-grammar.

Unless the criteria for seeing had been established it would be impossible to identify eyes as the organs of sight. P-grammar criteria are taxonomically prior to O-grammar criteria. In short no account of enabling conditions is possible without an account of the phenomena that they enable.

This point has been made often enough in different terms, for instance by the use of the phrase 'top/down'. However a robust defence is still required in the face of reductionist arguments from Davidson (1980) and reductionist *obiter dicta* from Crick (1994).

BIBLIOGRAPHICAL REFERENCES

Crick, Francis. *The astonishing hypothesis: The scientific search for the soul*. New York: Scribner, 1994. Print.

Davidson, Donald. *Essays on actions and events*. Oxford: Clarendon Press, 1980. Print.

Gibson, James J. *The senses considered as perceptual systems*. New York: Houghton Mifflin, 1966. Print.

Hart, H. L. A. *Law, liberty and morality*. Stanford, CA: Stanford University Press, 1963. Print.

Robinson, Daniel N. and Rom Harré. "On the primacy of duties". *Philosophy*. Oct. 1995: 513-532. Print.

Shaver, Kelly G. *The attribution of blame*. New York: Springer Verlag, 1985. Print.

Strawson, Peter Frederick. *Individuals*. London: Methuen, 1964. Print.

Vygotsky, Lev S. *Thought and language*. Cambridge, MA: MIT Press, 1962. Print.

Watkins, Mary. *Invisible guests: The development of imaginal dialogues*. Hillsdale, NJ: Analytic Press, 1996. Print.

Wittgenstein, Ludwig. *Philosophical investigations*. Oxford: Blackwell, 1953. Print.

CONSCIOUS AWARENESS AND BEHAVIOUR: WHAT DISTINGUISHES CONSCIOUS EXPERIENCE FROM UNCONSCIOUS PROCESSES?

CONCIENCIA CONSCIENTE Y COMPORTAMIENTO: ¿QUÉ DISTINGUE A LA EXPERIENCIA CONSCIENTE DE LOS PROCESOS INCONSCIENTES?

JOHN PEACOCK

University of Antwerp, Bélgica. john.peacock@ua.ac.be

RECIBIDO EL 13 DE MAYO DE 2012 Y APROBADO EL 24 DE MAYO DE 2012

RESUMEN ABSTRACT

El presente artículo analiza el concepto de conciencia consciente. La actividad del cerebro se puede dividir generalmente en dos categorías: procesos inconscientes y aquellos que contribuyen a la construcción de la experiencia consciente. Aquello de lo que somos conscientes y experimentamos y lo que no.

La comprensión de la conciencia requiere la comprensión de cuáles criterios separan a ambos. Por ello planteo un rol importante del comportamiento en esta distinción. Sugiero un medio para separar la experiencia consciente considerando el tipo de comportamiento expresable al que corresponden los procesos inconscientes.

This paper considers the concept of conscious awareness. The activity of the brain can generally be divided into two categories: unconscious processes and those that contribute to building conscious experience. That which we are consciously aware of and experiencing; and that which we are not.

An understanding of consciousness requires an understanding of what criteria separate the two. I argue a role for behaviour in the distinction. I suggest a means of separating out conscious experience by considering the sort of expressible behaviour that unconscious processes correspond to.

PALABRAS CLAVE KEY WORDS

Comportamiento, cerebro, conciencia consciente, experiencia consciente, conciencia, procesos inconscientes.

Behaviour, brain, conscious awareness, conscious experience, consciousness, unconscious processes.

Introduction

It is generally accepted that our conscious experience is a product of neural processes in the brain. If the brain were damaged in certain ways we would lose the ability to have certain conscious experiences, and if it stopped functioning it is generally believed that our conscious experience stops with it. But despite knowing them to be so intimately connected, how we go about explaining the connection between brain function and conscious experience seems to remain as much up for debate as ever.

Observation shows us one hundred billion interconnected neurons, and electrical impulses passing between them. We see the firing of these neurons leading to physical changes in a person, such as the stimulation of their muscles or the secretion of hormones etc. And we know that, somehow, out of this, also come our conscious experience¹.

I use the term conscious experience in the same way as Chalmers (*The problem*) in a recently reprinted article in this journal. Conscious experience is the 'hard problem' of consciousness. We are likely still a very long way from any satisfying account of this phantasm of experience, Nagel's 'what it is like'. And in this paper I do not intend to offer any solution. I have no blueprints to help us build a bridge across Levine's 'explanatory gap'.

Instead, the concept this paper will consider is the (easier?) problem of conscious awareness. The two ideas, conscious experience and conscious awareness are intertwined. What we are aware of is what we consciously experience; it is unintelligible to suggest that one has a conscious experience that one is unaware of. The common understanding is that, out of the myriad of electrical signals in the human brain, there are some that give rise to conscious experience and others that remain unconscious processes. In considering the structure of conscious awareness, what we want to consider is by what means we distinguish the two. Is there a means of saying why *this* pattern of electrical signals should equate to *this* conscious experience? While *that* pattern, equates to *that* experience?

¹ We *know* that it comes 'out of' the brain only in the sense that damage to the brain can demonstrably disrupt a person's consciousness, and that changes in states of consciousness, such as with a person entering sleep, are always accompanied with distinct changes in neural processes. However, it remains possible, as many have proposed that the consciousness itself is a product on more than the neural function alone.

More specifically I wish to argue for the role of behaviour (which will be defined more fully as we progress) in helping shape our awareness and so our conscious experience. What will be suggested is that conscious awareness only makes sense *in terms of* behaviour. Doing so, I aim to show, is the simplest way to reify mental states in bodily function.

The argument proceeds as follows: We first begin with a statement of the problem, and address the fundamental aspect that existing concepts fail to explain. I then consider the reasons why behaviour, as it will be defined, should be looked to to fill the gap, and in what sense I think we should consider it to do so. I then sketch a description of the manner in which behaviour could be said to separate the unconscious processes from conscious experiences. Given the intimate relationship between the structure of awareness and the substance² of conscious experience, I will end by briefly considering what implications this role for behaviour might have for any understanding of conscious experience itself.

Attention and its allies

Let's consider a person's response to a visual stimulus. Imagine there is some object in a room which is positioned such that light striking its surface is reflected towards a person's eyes. The light enters their eyes and stimulates the cells of their retina. This causes electrical signals to be sent to the brain which goes on to cause other neurons to fire, or prevent them from firing etc. This is something we know to occur.

In this situation we can say that a person's brain has received information about the object. However, whether or not that information leads to a conscious experience of the object, or otherwise, is not so easy to say. The mere existence of activity caused by a specific object does not guarantee that a person will possess a conscious experience of that object. Many sources are stimulating our senses continuously, yet we are not aware of all these sources, let alone the fact that certain

² I dislike using the term substance, as it implies that conscious experience is comprised of some physical 'stuff'. This is something I would argue is false. But I use it since most alternatives are no better. We are so far from an understanding, or are so completely misunderstanding the problem, that most of our terminology seems inadequate. Once again I direct the reader to Chalmers (*The problem*) account of the problem.

senses are being stimulated at all. The activity of neurons might lead to physical changes in a person, such as the regulation of their internal organs or the activity of their muscles. Yet even if an object were to cause such changes it might still only be classed as a subliminal stimulus: causing activity without any awareness of the object itself (e.g. Bornstein and Pittman, *Perception without*).

On top of that, the person might also remain oblivious to the physical change itself: one can perform complex actions like driving a car, or playing the piano, without any conscious awareness of the specifics of their movements.

Yet, out of all this activity does come a person's conscious experience. What is it about any neural activity that causes it to give rise to conscious experience? Most of the processes occurring in the brain are stimulated by things, and/or are causing changes in a person without any awareness. These are unconscious processes. What is puzzling is that these unconscious processes appear indistinguishable from those other processes involved in conscious experience. They are all the same electrical signals.

Conscious experience, it seems, is not connected with the nature of the signals themselves. Yet there must be some criteria by which they can be separated, something about certain neural activity, or what it does, that will tell us why it leads us to be conscious of specific things at specific times.

Since the stimulus that caused neural activity is not necessarily included in conscious experience and the physical changes that the brain causes are also not themselves necessarily included, it seems we must peer into the brain itself for a solution. When one is having a particular conscious experience, can we see something specific happening in the brain?

One currently popular way to look at this question is to try to look for neural correlates of consciousness (NCC): "*the minimal set of neuronal events and mechanisms jointly sufficient for a specific conscious percept*" (Koch, *The quest* 16)³.

³ Italics in the original.

It is thought, for example, that one needs to excite specific neural circuits in order to give rise to specific conscious experiences. The use of imaging technologies like fMRI lends support to this line of thought. With fMRI we are able to see that specific neural circuits do reliably 'light up' when a person reports possessing a particular conscious experience (Haynes and Rees, *Decoding mental*). We are led to believe that the activity of a particular neuron, or a specific group of neurons, is what constitutes a representation of a thing, and that the firing of these neurons is what leads to our conscious experience of that thing. But this use of NCCs raises a number of objections.

One objection can be found in Cohen and Dennett (*Consciousness cannot*), and their 'perfect experiment'. In brief: the experiment imagines that we separate, for example, the area that correlates with the perception of colour, yet still allow it to function. We imagine presenting the person with an object, a red apple say, and measuring their neural response and their outward behaviour.

Cohen and Dennett imagine the experimenter noticing the neural correlate of 'red' becoming 'active' when the person looks at the apple and concluding that the person has a conscious experience of the apple's redness. Yet, since the area in question is isolated and unable to cause any behavioural change in a person, the subject must emphatically deny that they see this redness in the apple. Who are we to believe?

If a person can't report, remember, make decisions about, plan to act in one way or another because of, and not realise they're having a particular experience when directly probed, then it hardly makes sense to say that they have that conscious experience, such as would be implied if we believed the experimenter with the apple.

Cohen and Dennett's position, therefore, is that it makes no sense to speak of a person having a conscious experience that they cannot access. In the absence of the functional capabilities mentioned above, there is no way we can verify any claim about a person having a conscious experience. Consciousness, they say, cannot be separated from function.

It is of course possible to suggest that the isolated group of neurons is itself having an independent conscious experience of red. And the

rest of the person's brain is having another experience. Individual phenomenal experience could be generated by the firing of specific NCCs e.g. Zeki (*Localization and*).

But this only raises further objections. Firstly, it is a claim that is impossible to verify scientifically (since we rely on a person's report to ascertain what they are experiencing consciously). But more fundamentally, we are left asking whether the group of neurons, before it was isolated, could have an independent conscious experience.

If we say it can have a conscious experience of red in isolation, that the firing of those neurons is what, somehow, 'creates the redness', then surely it and all the other areas of a person's brain can conjure conscious experience separately and independently.

Yet a person experiences a single coherent conscious experience, and we remain wondering how these separate experience producing areas build that. If the different areas of the brain can conjure the stuff of experience, the qualia, independently, it might imply that there is something central to which this is passed onto to create an individual's experience.

We can imagine adding a switch such that we can at once connect and disconnect the isolated brain region. Must we then accept that we can switch on and off a person's conscious experience of the red?

If we do, and accept that the group of neurons in isolation is what creates the particular qualia, we must explain how this qualia happens to be passed onto whatever it is passed onto. How does connecting a group of neurons that generate a qualia, to another group give coherence to the entities conscious experience?

If each brain region can produce and experience the qualia in isolation, why isn't that all they do? Why would connecting them together allow the qualias to be become interlinked and, seemingly, experienced together?⁴

⁴ The notion that we have a single 'stream of consciousness', a single thing that is constructed is, I suspect, entirely false. I do not believe that the brain constructs some sort of 'movie' and my statement here are not meant to endorse such a position. See Blackmore (*There is*), for a challenge to the 'streams of consciousness'.

We know that functions are localised in different brain regions, and that removing or damaging a specific region leads to specific impairment in a person's physical and mental abilities, including their ability to have particular conscious experiences.

Yet the best this can tell us is that the firing of a particular group of neurons is a necessary condition for a person being conscious of a particular sort of thing. But having those neurons fire cannot, on its own, account for our conscious experience. Or, at the very least, there is no convincing account of the how an isolated qualia theory would work.

The existence of any homunculus in the brain, to which information is passed and which is the conscious experiencer, has been widely refuted, notably by Dennett (*Consciousness*). If we do not wish to champion such a homunculus, and we choose to believe that individual elements of phenomenal consciousness are a result of specific, independent neural circuits, we must provide a means by which connecting them allows their experiences to seem shared.

Moreover, proposing that a person experiences a particular thing when that NCC is 'active', for example, does not really offer a satisfying explanation even for the NCC still entirely connected. What we still have is a group of neurons, and electrical signals firing between them. Not only is there still a huge gap between them firing and the conscious experience, the hard problem, we have not offered a way of separating these particular electrical signals from all the other 'unconscious' ones, the 'easy' problem of awareness. The firing of those neurons that make up the NCC may be important for a specific conscious experience, we can see the correlation, but as far as we can see they are doing the same as all the others.

We have no way of saying why they give *this* experience as opposed to *that*. And, as we've already established, any appeal to what stimulated the activity will not suffice. A satisfying account of awareness must tell us something of how we distinguish between neural signals.

If we stick to the NCC, and show that a particular pattern of neural activity is reliably active when a person is aware of *this*, we must go on to explain a connection between that particular pattern and *this*, which must include what is separating that pattern from all the other

neural activity that, presumably, is not currently building a conscious experience.

At the end of their paper, Cohen and Dennett conclude that:

A true scientific theory will say how functions such as attention, working memory and decision making interact and come together to form a conscious experience.
(*Consciousness* 362)

This appeal to functions such as attention is where I begin to differ from them. Attention, as an explanation, is no different from the NCCs described above. Attention is proposed as a specific function that selects one thing for us to be conscious of over another. It picks out of the multitude of neural signals what consciousness is to focus on. Attention, in other words, is supposed to account for awareness.

The concept of attention ties in well with ideas such as the 'global workspace' of Baars (*A cognitive, In the theatre, The global*), where consciousness results from unconscious activity in different brain regions becoming coherent, particularly with respect to frontoparietal and medial temporal cortices.

Or Dennett's (*Consciousness, Are we explaining*) own somewhat similar, idea of 'fame in the brain', or 'political clout'. In these it is suggested that activity in the brain becomes consciously available by harnessing certain cortical resources, or achieving a certain influence in the brain. These sorts of things are what many suggest are driving attention.

But even if we can see that global influence correlates with consciousness we are still left with the task of explaining how neural activity actually relates to the awareness we have of a specific thing. We are ultimately in the same situation as with the NCC.

Look again at the brain and all we see is electrical signals. How do we say they relate to a concept, or an idea? And when the signals spread and influence other areas of the brain, how do they still relate to the same concept? If there are competing representations in the brain we are left asking how the neural activity corresponds to the representation, and how that idea endures as electrical signals dissipate or spread. Then if one of these representations gains political clout, this does not

really tell us what differentiates this activity from other activity that is remains unconscious.

If we suggest that there is competing activity in the brain and that, for whatever reason, some happens to become dominant over large areas of the brain and this is what dictates our present conscious experience, we have to then ask what links that activity, with its global influence, to *that* specific conscious experience.

Proposing a thing is conscious because of particular global influence of activity denies us the 'perfect experiment' as an objection, since separating a particular region removes the ability to be influential globally. But as an explanation it is still lacking in a fundamental way. We still only have electrical signals, there is no way to associate them with a particular conscious experience, to say how they correspond to a representation of a particular thing, even after the pattern changes and dissipates.

And if we assume that the pattern of neural activity, somehow, does constitute a representation of the object, we still need to know what difference its global influence has on it that allows it to become conscious experience. Globally influential neural activity remains neural activity, why should it dictate what we are consciously aware of above all the other neural processes? The same challenge must also be offered to any proposal based on, say, gamma synchrony in the brain e.g. Hameroff (*The "conscious pilot"*) or similar ideas.

If we say that, by virtue of this process (gamma synchrony, global influence or whatever else) a thing becomes part of conscious experience, we have to have some way of associating that specific activity with that specific experience. We need a way to separate out neural activity between the conscious and unconscious.

Showing the correlation between the activity of some neural circuit or process and some specific conscious experience will not be enough; a satisfying account needs to give an explanation that goes beyond the mere experimental observations of their coincidence to say something about the nature of the neural signals or what they are doing.

If we propose attention is what separates neural activity into the consciously experienced from the unconscious, the thing that, say,

makes me aware of my heart beating at sometimes but generally leaves me blissfully unaware, we are left with explaining what attention is selecting between, as well as what exactly attention itself is.

If we think working memory is involved, we need to understand what that is, how exactly a group of neural signals equates to a memory. In what follows I propose to simplify this process of awareness by cutting out the middle men: attention and its allies.

Instead, I argue that those things can be completely understood in terms of behaviour, and that conscious awareness, from which a specific conscious experience results, can be readily constructed out of unconscious neural process by considering behaviour alone.

A role for behaviour

An immense number of ideas have been proposed in full, or part explanation of the various concepts surrounding consciousness. These range from the conventional to the bizarre, we find ideas that are spiritual and religious, and ideas that are wholly material and scientific. We might invoke string theory or quantum mechanics (like Penrose (*Shadows of*), or Stapp (*Mindful universe*)); some as yet unknown fundamental physics could be postulated, perhaps supporting a form of panpsychism like that of Skrbina (*Panpsychism in*).

We can look into the detail of how the brain operates in the hope of finding something unique. To provide an account of the structure of awareness, I propose to take the one thing we know the brain to do with absolute certainty: produce physical, behavioural changes in a person.

Before beginning my own account, a number of clarifications are necessary. My use of the term behaviour, as already hinted, refers to all the physical changes in a person that result from neural activity. This is the aforementioned stimulation of muscles, regulation of internal organs, secretion of hormones etc. All of these things work to alter what an organism does in its present environment and so we refer to them as behavioural changes.

Secondly, I am not going to suggest that the actual performance of behaviour is necessary for consciousness. Any interdependence

between the actual physical bodily change of a behaviour and consciousness is demonstrably false. People with locked in syndrome, for example, can be conscious of things without the ability to physically enact behaviours (see e.g. Owen et al. *Detecting awareness*).

What is going to be suggested, however, is that any conscious experience can only sensibly be considered in terms of behaviour. Imagining an action, for example, excites the same neural circuitry as its actual performance (e.g. Kuitz-Buschbeck et al. *Effector-independent*).

Meaning that neural activity, presently and subsequently occurring can still be influenced by it *as if* the behavioural change did occur. Whether or not behaviour is physically enacted needn't matter, neural activity can be said to correspond to certain behavioural changes even without them occurring. The activity associated with an 'internal thought' of an act has been experimentally shown to be the same as the acts actual performance. Its correspondence to behavioural changes is what, I intend to argue, separates neural activity into conscious experiences and unconscious processes.

Seeing that activity is occurring, we don't simply consider that it is electrical signals; we consider it in terms of certain physical changes it corresponds to. We might see the necessity of, for example, the activity of certain grid cells in allowing a person to 'know' their location in a particular space, and so have conscious awareness of their being in that particular position (see Moser, *Place cells*). But it is not simply the firing of the neurons that 'is' that awareness, the awareness comes from the fact that the firing of that particular pattern of neurons corresponds to certain behavioural changes, or directly promotes other activity that corresponds to those changes, and it is that 'behaviour' that constitutes awareness.

There are two reasons we can argue for the role of behaviour in conscious awareness: firstly is the consequences of removing the ability to perform certain behaviours (or removing the fact that certain neural activity corresponds to them), and second is the consequences of having a person perform a specific behaviour (or having a person's neural activity correspond to certain behaviours). We will consider them in turn.

First, like in the perfect experiment previously described, we can ask whether stripping away the ability of a person to express that they

are having some experience (or some behaviour related to having the experience) implies that they cannot have it. Consider the relationship between pain and pain behaviour⁵. Is it possible for a person to be in pain without any related behaviour?

Imagine that a person can have pain experiences whilst their neural activity does not, and in future cannot, correspond to any sort of behaviour that relates to that pain. Pain might be like a disease and the behaviour the symptoms.

Without any behaviours associated with pain, we would have to imagine that the person we are observing conducts their life without the ability to ever display any behaviour related to that pain. One could not speak of it, be hindered in their movement by it, decide to act because of it etc. But to have an experience of pain, a person must also be aware that they have pain (this to me seems to be true by definition: what we are conscious of, and what forms part of our conscious experience, is what we are aware of). But to say that I am aware that I am in pain, yet have this awareness unable to influence any possible behaviour, seems to be a contradiction.

If there is no possibility for the neural activity that is 'causing' the conscious experience of pain to correspond to, or cause, certain specific physical changes related to that experience, then to suggest that a person still has that experience leaves one in the same situation as the experimenter and his apple.

Would any sensible experimenter claim that a person is in pain, after measuring their neural activity, whilst the person themselves carries out all the activity of their lives without any possibility of displaying behaviour related to this pain? Including the ability to say that they are in pain when asked? (And this is for a person otherwise perfectly capable of expressing themselves through speech). The disease can be separate from the symptoms in the sense that a person can have a disease without awareness of it. But with phenomenal consciousness, surely the experience and awareness must be tied?

If one has a disease that affects them only subtly, such that it influences what they do, yet leaves them unaware of the changes, then it must

⁵ A thorough examination of this subject is provided in Hickman (*Is it Intelligible*). Her conclusion is, like mine, that they cannot be separated.

be thought of like a subliminal stimulus, and one is hardly conscious of that. To have a conscious experience, such as pain, subliminally is senseless. And to suggest one has a full-blown conscious experience, such as pain, without the possibility for this to alter *any* behaviour is also senseless.

Any behaviour includes the ability to report, remember, make decisions about, plan to act in one way or another because of, and realise you're having a particular experience when directly probed. Minus all of these things, one must be left with a very bizarre view of experience. Awareness implies the ability to behave in a particular way, and conscious experience implies awareness.

A similar argument can be presented from the opposite direction: that if a person, were able to express certain behaviours then they would necessarily be aware of certain things. If the person in the perfect experiment said "yes I see the apples redness", if they reported (with sincerity) its colour to be red, if they would later recall having seen the apples redness and could identify this colour from a chart, if they were able to distinguish its colour from others and say it is the same as another 'red', if they made use of its colour to judge its edibility etc. then it hardly seems fair to suggest that they have no conscious awareness of it (and, also assuming that they are not a philosophical zombie, this implies that they would also have a conscious experience of it).

If I can verbalise the existence of the object in the room, then I must necessarily be aware, to some extent, of its existence in the room. Which of its properties I am aware of is unclear, but if I speak of it explicitly one doesn't imagine that I have no real conscious experience of it. It can no longer be a subliminal stimulus, affecting me unconsciously, if I display certain behaviours.

Thus if we take away all possible behaviours, it is difficult to argue that a person is aware of a thing and so having a conscious experience of it, and if we introduce specific behaviours we are left in a position where we must conclude that a person does have awareness and thus experience. Having your behaviour influenced by a thing does not imply your awareness of it, but simply displaying certain specific behaviours, in relation to a thing, does. From here we can suggest something about the role of behaviour in constructing awareness.

Conscious experience from unconscious processes

Rather than look to promote something like a specific process of 'attention' as choosing what we are consciously aware of, and so experiencing, it is simpler to discard it and see the process entirely in terms of behaviour. Let us take an example. A person possesses a body and is at times aware of that body. A person is never, however, aware of their entire body. My attention, at times, may be focussed on my hand (I am aware of where it is and what it is doing), but at the same time I am not aware of my foot, although my foot may be doing things (like helping me walk down the street).

At another time my focus is shifted and I am aware of my foot and not my hand. The question is what causes my awareness at different times, or more fundamentally: what is my awareness.

Previously I argued that it makes little sense for a person to be consciously aware of a specific object yet for their neural activity never to allow any form of behaviour that would demonstrate that awareness. So if the mind, and conscious awareness, is a product of neural activity, and neural activity must *correspond* to certain behaviours for a person to have awareness of a specific thing, can't we suggest that it is the fact of having activity that corresponds to specific behaviours that accounts for our awareness of a specific thing?

In other words, I would not say I am aware of the position of my hand because my brain has constructed a model of my body and its position in space, and because 'attention' has selectively decided to make me aware of it rather than my foot. I say that I am aware of the position of my hand because my neural activity happens at that moment to correspond to behaviours that constitute awareness of my hands position⁶.

When I am consciously aware that my hand is resting on the arm of my chair, and not on the desk in front of me, it is because my neural activity corresponds to sets of behaviours that would be appropriate to its being in that location. (Once again I must stress that the actual

⁶ What, specifically, those behaviours need to be is a question I leave open. I have argued that in the absence of certain behavioural abilities (or a brain that can have neural activity that corresponds to such abilities) it seems that one cannot be conscious of certain things, but identifying specifics is a subject that would require much more detailed considerations than this paper allows.

performance of the behaviours would not be necessary). In other words we could say that the neural activity that constitutes my awareness of the hands position, so constitutes this, and not some other thing, by virtue of their relation to those particular expressible behaviours.

In the famous rubber hand illusion, a person is tricked into experiencing their hand (that is all sensations related to it) as occurring in the same position as a rubber hand placed nearby, rather than coincident with the location of their real hand (Botvinick and Cohen, *Rubber hand*).

One can even trick a person into experiencing any random object as being their hand! One could suggest that one's brain has constructed the model and projects their experience of their hand to a particular point in space, and that the experiment tricks it into projecting the experience elsewhere. To my mind this explanation is overly complicated.

If we take a knife and attempt to stab the rubber hand, a person will show the same physiological response that they would if their real hand was so attacked, a response that does not occur when some unrelated object is threatened (Armel and Ramachandran, *Projecting sensations*).

But if they did not exhibit such behaviour, if their neural activity did not correspond to any behavioural changes in keeping with their hand being in the position of the rubber one, then we could not sensibly say that they experience it in that place. They themselves would not know that their hand was in a particular position if they couldn't possibly behaviourally respond to its being in that position. If my neural activity, on observing the rubber hand, entirely corresponded to the same range of behaviours as it does to another inanimate object, then I couldn't possibly also experience that as my hand.

Our awareness of the position of ourselves in relation to other objects, and of the relative positions of other objects, is necessarily accompanied by behaviour in keeping with its being so. We are only aware, and can experience, a thing as 'over there' if we can and/or do behave in appropriate ways to its being so.

The simpler explanation, again, takes the fact that our neural activity happens to correspond to certain behavioural changes as being central to what we are aware of and thus what we consciously experience.

As a further example, the reason I am generally unaware of my heart beating or my breathing, even though they are products of neural activity just like what I am conscious of, is because the behavioural change itself is what helps build my conscious experience.

I am not aware of my arm because I move my arm, I am aware of it when my behaviour constitutes awareness of this, which may only partly be a result of the specific behaviour that is moving the arm.

The neural activity that is controlling my heartbeat may contribute to my awareness of 'something'⁷, but the fact that it causes *this* behavioural change needn't give it any connection to my awareness of *this*. Certainly, my heart needs to beat for me to be aware of it! But its beating doesn't constitute this awareness.

At certain, other, times I am able to count the beats and verbalise this or conduct other movements in time with its beating. It is then that I might be said to be aware of its beating, and I am so aware because I can, at that time, enact those behaviours. The extent of my awareness, the detail of the awareness, is dictated by the specific behaviours that my cognitive faculties manage to produce.

In the original example of an object in the room (with light reflecting from it into a person's eye) and given the scheme I have sketched out, we no longer need to ask anything of the brain other than that it processes those electrical signals, and that those correspond to certain physical changes in a person: something we know it is already doing. There is nothing extra to add, it just corresponds to behavioural changes.

Neural activity causes us to act in different ways and be aware of different things, and the nature of what we are specifically aware of is dependent on the specifics of what behavioural changes our present neural activity constitutes. We suddenly become aware of the object in the room, or of our hand, or of ourselves, because, given whatever stimuli we have received, and given the ways our own brain processes those signals, we end up with activity corresponding to behaviour that

⁷ Changes in heart rate do play a role in, for example, the regulation of a person's emotions (e.g. Kalat and Shiota *Emotion*). And our emotional state is something which we are generally aware of, or contributes to our awareness of some other thing. Thus, this particular behaviour, the heart rate, could be said to help construct our awareness of something via its role in constructing an emotional state.

constitutes (or perhaps a better phrase would be 'is in keeping with') awareness of 'this' and not 'that'.

Summary and conclusions

In this paper I have attempted a brief sketch of how we could understand the dependence of conscious awareness with behaviour, the physical changes in a person that neural activity corresponds to. This, I believe, is the simplest way to reify conscious awareness in bodily function. With this description we do not ask anything more of the brain than what we know it to do with certainty.

We know that information about the world is encoded into electrical signals. These are passed around the brain, stimulating patterns of activity in the neurons. It is generally accepted that our conscious experience is a result of this activity, which the brain somehow represents a thing in its activity and thus leads to what we experience.

My suggestion is that we understand this representation in terms of behaviour. I suggest that we look to what behaviours neural activity corresponds to in order to account for awareness, and so distinguish between unconscious processes and conscious experience.

In other words, awareness is to be understood in terms of behaviour. We represent the world based on expressible behaviour (even if we cannot, or do not, actually express it). Now, instead of asking whether or not conscious thoughts cause behaviour (like Baumeister, Masicampo and Vohs, *Do conscious*), we can suggest that conscious thought is, effectively, constructed out of behaviour.

A fuller development of the relationship between specific behaviours and awareness of specific things needs further work. But, if we think in terms of expressible behaviour, we can approach various aspects of consciousness by asking what disallowing certain behaviours will do to what we can be aware of.

We might, for example, attempt to give an account of our awareness of ourselves by asking what behaviours we would need to take away before we can no longer sensibly say we are self-aware. For example, if we remove certain language abilities, would it still make sense to be

aware of oneself? An account in terms of behaviour, if correct, provides us with a convenient means of approaching these various aspects.

However, this description is only a sketch. If correct, we can show that a person's conscious awareness only makes sense in terms of behaviour, but we must acknowledge that behaviour itself does not provide an explanation for the substance of conscious experience.

The experience and the awareness coincide, yet seem to be so separate. We might bring out our philosophical zombie, with all the behaviours but without the experience, as an example of that (Chalmers, *The conscious*).

Yet we can also see that the experience that we do have *is* tied to behaviour. Perhaps this realisation can provide some help in constructing a satisfying account of conscious experience, or perhaps that explanation will forever remain inaccessible.

A non-reductive explanation may be called for (as, again, with Chalmers (*The problem*)), but if we consider my account correct we must factor in why experience would be dependent on neural activity's relationship with expressible behaviour.

BIBLIOGRAPHICAL REFERENCES

Armel, Kathleen Carrie and Vilayanur Ramachandran. "Projecting sensations to external objects: Evidence from skin conductance response". *Proceedings Royal Society B-Biological Sciences*. Jan. 2003: 1499-1506. Print.

Baars, Bernard J. *A cognitive theory of consciousness*. New York: Cambridge University Press, 1988. Print.

---. *In the theatre of consciousness: The workspace of the mind*. New York: Oxford University Press, 1997. Print.

---. "The global brainweb: An update on global workspace theory". *Science and Consciousness Review*. Oct. 2003. Online.

Baumeister, Roy F, Masicampo, E. J. and Kathleen D. Vohs. "Do conscious thoughts cause behavior?" *Annual Review of Psychology*. Nov. 2010: 331-361. Print.

Blackmore, Susan. "There is no stream of consciousness". *Journal of Consciousness Studies*. May. 2002. Print.

Bornstein, Robert F. and Thane S. Pittman, ed. *Perception without awareness: Cognitive, clinical, and social perspectives*. New York: Guilford Press, 1992. Print.

Botvinick, Matthew and Jonathan D. Cohen. "Rubber hand 'feels' what eyes see". *Nature*. Feb. 1998: 756. Print.

Chalmers, David. *The conscious mind: In search of a fundamental theory*. New York: Oxford University Press, 1996. Print.

---. "The problem of consciousness". *Discusiones Filosóficas*. Jul.-Dic. 2011: 29-59. Impreso.

Cohen, Michael A. and Daniel C. Dennett. "Consciousness cannot be separated from function". *Trends in Cognitive Science*. Aug. 2011: 358-364. Print.

---. "Response to Fahrenfort and Lamme: defining reportability, accessibility and sufficiency in conscious awareness". *Trends in Cognitive Science*. Jan. 2012: 139-140. Print.

Dennett, Daniel C. *Consciousness explained*. London: Allen Lane-Penguin Books, 1992. Print.

---. "Are we explaining consciousness yet?" *Cognition*. Apr. 2001: 221-237. Print.

Hameroff, Stuart. "The "conscious pilot"-dendritic synchrony moves through the brain to mediate consciousness". *Journal Biological Physics*. Jan. 2010: 71-93. Print.

Haynes, John-Dylan and Geraint Rees. "Decoding mental states from brain activity in humans". *Nature Reviews Neuroscience*. Jul. 2006: 523-534. Print.

Hickman, Waights Natalie. "Is it Intelligible that an Organism with no Pain-Behaviour should be in Pain?" *Journal of Consciousness Studies*. 2011: 67-93. Print.

Kalat, James W. and Michelle N. Shiota. *Emotion*. Belmont: Wadsworth Publishing, 2006. Print.

Koch, Christof. *The quest for consciousness: A neurobiological approach*. Colorado: Roberts & Company Publishers, 2004. Print.

Kuhtz-Buschbeck, Johann P., et al. "Effector-independent representations of simple and complex imagined finger movements: a combined fMRI and TMS study". *European Journal Neuroscience*. Dec. 2003: 3375-3387. Print.

Lamme, Victor A. F. "Towards a true neural stance on consciousness". *Trends in Cognitive Sciences*. Nov. 2006: 494-501. Print.

Levine, Joseph. "Materialism and qualia: The explanatory gap". *Pacific Philosophical Quarterly*. 1983: 354-361. Print.

Moser, Edvard, et al. "Place cells, grid cells and the brain's spatial representation system". *Annual Review of Neuroscience*. Jul. 2008: 69-89. Print.

Nagel, Thomas. "What is it like to be a bat?" *Philosophical Review*. Oct. 1974: 435-450. Print.

Owen, Adrian M., et al. "Detecting awareness in the vegetative state". *Science*. Sep. 2006: 1402. Print.

Penrose, Roger. *Shadows of the mind*. New York: Oxford University Press, 1996. Print.

Skrbina, David. *Panpsychism in the west*. Boston: MIT Press, 2007. Print.

Soon, Chun Siong, et al. "Unconscious determinants of free decisions in the human brain". *Nature Neuroscience*. May. 2008: 543-545.

Stapp, Henry. *Mindful universe: Quantum mechanics and the participating observer*. New York: Springer-Verlag, 2007. Print.

Zeki, Semir. "Localization and globalization in conscious vision". *Annual Review of Neuroscience*. 2001: 57-86. Print.

THE POVERTY OF ECONOMISM: FREEDOM, CALCULATION, AND THE LAW

LA POBREZA DEL ECONOMICISMO: LIBERTAD, CÁLCULO Y LEY

MARK A. NOTTURNO

Interactivity Foundation, USA. notturmo@interactivityfoundation.org

RECIBIDO EL 26 DE ABRIL DE 2012 Y APROBADO EL 28 DE MAYO DE 2012

RESUMEN ABSTRACT

Karl Popper y Friedrich von Hayek son recordados como dos de los mayores defensores en el siglo XX, del concepto de sociedad abierta. Ambos nacieron en Viena. Ambos se sintieron atraídos hacia el socialismo en sus años mozos, y los dos escribieron críticas seminales contra el socialismo, que revelaban sus defectos fundamentales. También eran muy buenos amigos, que se ayudaron mutuamente en sus carreras, y generalmente se los consideró aliados filosóficos. Sin embargo, los puntos de vista de Hayek sobre la democracia, la racionalidad y el economicismo, eran fundamentalmente opuestos a los de Popper -e incluso quizás- a la sociedad abierta en sí.

El presente artículo, se centra en sus diferencias sobre el economicismo. Argumenta que la crítica popperiana al economicismo de Marx, también puede aplicarse a Hayek; que éste estaba "preparado" para aceptar el socialismo si pudiera ser tan eficiente y productivo como el mercado; que Popper no; y que es imposible realizar la idea de Hayek sobre la libertad por la misma razón que éste pensaba que el socialismo no podría tener éxito.

Karl Popper and Friedrich von Hayek are remembered as two of the 20th century's greatest proponents of open society. They both were born in Vienna. They both were attracted to socialism early in life. And they both wrote seminal critiques of socialism that revealed its fundamental flaws. They were also very close friends, helped each other in their careers, and were generally regarded as philosophical allies. But Hayek's views about democracy, rationality, and economicism are fundamentally at odds with Popper's -and perhaps even- with open society itself.

This paper focuses upon their differences about economicism. It argues that Popper's critique of Marx's economicism also applies to Hayek; that Hayek was 'prepared' to accept socialism if it could be as efficient and productive as the market; that Popper wasn't; and that it is impossible to realize Hayek's idea of freedom for the very same reason that Hayek thought socialism could not succeed.

PALABRAS CLAVE KEY WORDS

Problema del cálculo, economicismo, Hayek, ingeniería paso a paso, Popper, sociedad abierta, imperio de la ley.

Calculation problem, economicism, Hayek, piecemeal engineering, Popper, open society, rule of law.

"A policy of framing institutions to safeguard democracy must always proceed on the assumption that there may be anti-democratic tendencies latent among the ruled as well as among the rulers" (Karl R. Popper).

Introduction

Karl Popper and Friedrich von Hayek were two of the twentieth century's greatest proponents of open society—and their lives and intellectual careers, which spanned nearly the entire century, were similar in many ways. Both were born at the turn of the century in Franz Josef's Vienna. Both studied at the University of Vienna and were attracted to socialism as young men. Both immigrated to England and taught at the London School of Economics. Both wrote controversial books (Hayek (1944) Popper (1945)) that defended democracy and pointed out the unintended totalitarian consequences of socialism. And both of these books, published just one year apart from each other, attracted wide audiences and are still selling well nearly sixty years after they first appeared. These facts are not coincidental. Popper, though only three years his junior, always regarded Hayek as a 'father-figure'¹. He presented his critique of historicism first in Hayek's seminar in London. And Hayek, the first of the two to establish an international reputation, played an important role in finding publishers for Popper's work², in promoting Popper's ideas about science and open society, and in bringing Popper himself to the London School of Economics.

Popper felt that Hayek had saved his life by helping him to find a job and a publisher for his book (Popper, *Unended quest* 120). He included references to Hayek in many of his books and dedicated *Conjectures and refutations* to him. Hayek, for his part, also referred to Popper in his work. He tried for years to raise money to fund an institute to promote their work. And he told interviewers, late in life, that he and Popper were very close friends and that "to a very large extent I have agreed with him, although not always immediately",—adding that "on the whole I agree with him more than with anybody else on philosophical matters" (Kresge & Wenar 51).

¹ Popper, in writing this, quickly added that he did not mean 'father-figure' in the Freudian sense.

² Popper's *Poverty of historicism* was first published in the journal "Economica", which Hayek edited; and Hayek went to great pains to find a publisher for Popper's *The open society and its enemies*. Popper, insofar as this is concerned, wrote in its 'Acknowledgements' that the book would not have been published without Hayek's interest and support.

The road to serfdom and *The open society and its enemies* were vilified by 'left' leaning American academics when they first appeared³. That was over half a century ago. The war against socialism has since been won in the United States—at least for the moment—and 'open society' is now on the rise in many of the countries of the former Soviet Union and socialist bloc.

Hayek's ideas are now taught in many American universities, and several libertarian 'think tanks' in the Washington D.C. area proudly claim him as an intellectual founder. It is unfortunate, however, that these institutions do not pay more attention to Popper. I say this because Popper and Hayek differed in subtle ways—and in ways that are perhaps not so subtle—about democracy, rationality, and economism; because these issues are, or should be, fundamental to the libertarian cause; and because Popper's views about them are, in my view, decidedly preferable.

Hayek and Popper agreed that rationality has its limits, that authoritarian philosophies have distorted its meaning, and that the idea that social science can predict the course of history betrays a fundamental misunderstanding of the scientific method. They also agreed that democracy is not an end in itself but a means toward safeguarding freedom—and that it is not a foolproof means toward this end, but one that is easily manipulated by totalitarians. And they agreed, at a time when the idea did not seem as clear as it does today, that the market is more productive and efficient than central planning—and that socialism leads to fascism and totalitarianism, despite the good intentions of its proponents. These agreements mark broad strokes in the history of political and economic thought, and I hope that nothing I say in the course of this paper will be misinterpreted as denying this fact. Writ large upon the history of philosophy, Popper and Hayek have fought many of the same battles against many of the same enemies. But if I am right, then Hayek's views about democracy, rationality, and economism are fundamentally at odds with Popper's—and, perhaps, even—with open society itself.

³ Hayek, in his preface to the 1976 edition of the book, wrote that "just to indicate the character of a widespread reaction, I will mention merely that one well-known philosopher, who shall be nameless, wrote to another to reproach him for having lauded this scandalous book, which 'of course [he] had not read!'" The philosopher in question was Rudolf Carnap, then at the University of Chicago. The philosopher to whom Carnap wrote was Karl Popper. (See, in this connection, my 'Popper's Critique of Scientific Socialism, or Carnap and His Co-Workers' in my *Science and the open society: The future of Karl Popper's philosophy*).

Popper and Hayek differed about whether and to what extent our societies are well served by deliberate attempts at intervention. They also differed as to whether our democracies are better served by institutions that are designed to help us elect the best leaders, or by institutions designed to protect us against the leaders we elect. And they differed, perhaps most importantly, as to whether we should value freedom as a means to prosperity or prosperity as a means to freedom. These issues lie at the heart of liberal thought. And they have far reaching consequences. Indeed, Hayek, if I am right, designed reforms to our electoral system that would actually transform a democracy into what Popper would have regarded as a tyranny. I cannot do justice to these differences in this short paper. But I want, in what follows, to focus upon Hayek's economism, and upon his idea of freedom. If I am right, then Hayek, contrary to Popper, would have been prepared to accept totalitarianism if it could be shown to be as efficient and productive as the market. And if I am right, then it is impossible to realize Hayek's idea of freedom for the very same reason that he thought socialism cannot succeed.

Tyranny and the democratic state

I suspect that the idea that Hayek proposed reforms to our electoral system that would transform a democracy into what Popper would have regarded as a tyranny may sound strange to many readers. But it illustrates the differences between Popper and Hayek so well that I want to briefly explain it before turning to Hayek's economism.

Popper thought that Plato had set western political philosophy off on the wrong foot by making the question 'Who should rule?' its primary problem. He said that it is nearly impossible, once we pose this question, not to offer some utopian answer, such as 'the best' or 'the wisest' — and that the more pressing political problem, in any event, is how to get rid of our rulers when they turn out to be not as good or as wise as we had hoped them to be. Popper thought that we need to distinguish between two and only two forms of government insofar as this is concerned: between democracies, which possess institutions, such as elections, that allow us to dismiss our rulers without violent revolution — and tyrannies, which do not (Popper, *The open society* 161). And he argued that 'the criterion of a democracy' is whether or not "the rulers—that is to say, the government— can be dismissed by the ruled without bloodshed" (Ibid. 161).

Hayek, on the other hand, thought that our democracies are compromised by a felt need on the part of elected officials to pander to the desires of their constituencies in order to ensure their own reelections. He thought that we could solve this problem by electing our representatives for a period of fifteen years, without possibility of reelection. And he proposed that our electoral conventions should be revised so that legislatures would be composed of representatives elected, in effect, for life by members of their own age groups, who would vote once and only once in their lifetimes at the age of forty-five (Hayek, *New studies* 160-1).

Hayek characterized his proposal as 'utopian'; and there are at least two utopian features that sharply distinguish it from the way in which Popper thought about democracy. The first is that it is motivated almost entirely by the problem 'Who should Rule?' This might not be obvious at first glance, since the proposal is an attempt to solve the problem of officials pandering to their electorates. But Hayek's idea would not even be a tentative solution to this problem, were it not for his assumption that an electorate voting at the age of forty-five, and voting only for candidates in its own age group, is more likely than otherwise to elect the best and wisest leaders.

Hayek, indeed, argued for this assumption by saying that people are more likely to know who is best qualified to lead within their own age group, and that people at the age of forty-five are both experienced enough and mature enough to make such judgments. This preoccupation with finding a way to get the best rulers may seem innocuous. But it leads directly to a second utopian feature of Hayek's proposal that is far more dangerous. For if we can really design a system that enables us to get the best and wisest leaders, then there would be no reason whatsoever to think about how to get rid of them if something goes wrong. And Hayek's proposal, in any event, leaves us with no institutional method for getting rid of them. For the members of the electorate in Hayek's system would vote once and only once in their lifetimes. This means that they would have an opportunity to vote their rulers into office, but no opportunity to vote them out.

But Popper, as we have seen, thought that democracies are distinguished from tyrannies by the existence of institutions that enable us to get rid of our rulers without violence and bloodshed. And this means that Hayek's proposal, in lieu of other institutional arrangements that would enable us to get rid of our elected officials (and Hayek, so far as I can see, offers none), would transform a democracy into a system that Popper would regard as a tyranny.

I think that it is reasonably clear that Hayek intended his proposal to yield an improvement of democracy. And I do not know whether or to what extent Hayek accepted Popper's criterion of democracy. But this is not the only issue about which they disagreed. And Hayek's economism seems to imply a willingness to accept tyranny on his own terms.

The poverty of economism⁴

Popper criticized Marx for his *historicism*, which he described as the belief that the course of history is predetermined by scientific laws. But he described Marx's historicism as 'economism', since:

Marx, in opposition to Hegel, contended that the clue to history, even to the history of ideas, is to be found in the development of the relations between man and his natural environment, the material world; that is to say, in his economic life, and not in his spiritual life. (Popper, *The open society* 104)

I do not know whether Marx ever used the term 'economism' himself⁵. But Popper used it to describe,

the claim that the economic organization of society, the organization of our exchange of matter with nature, is fundamental for all social institutions and especially for their historical development. (Ibid. 106)

He said that economism "is perfectly sound, so long as we take the term 'fundamental' in an ordinary vague sense, not laying too much stress upon it" (Ibid. 106)⁶. But he criticized Marx for overemphasizing it, and for trying to reduce all thoughts and ideas to economic conditions (Ibid. 107). He said that "the general importance of Marx's economism can hardly be overrated", but that "it is very easy to overrate the importance of the

⁴ 'The poverty of economism' alludes to Popper's *The poverty of historicism*, which alludes to Marx's *The poverty of philosophy*, which alludes to Proudhon's *The philosophy of poverty*. I use it to underscore the fact that Popper thought that Marx's historicism took the form of economism.

⁵ Lenin, who did use the term, criticized the Marxist school of 'economism'. But he had something very different in mind: the view, namely, that socialist revolution in underdeveloped Russia was premature and that Russian Marxists should restrict their activities to helping workers secure better economic conditions. The term has a derogatory connotation — so that thinkers are less likely to declare themselves proponents of economism than they are to be accused of it.

⁶ Italics by the author.

economic conditions in any particular case” (Popper, *The open society* 107). It is important, however, to understand that economism is not peculiar to Marx, and that it is entirely consistent with a market approach to economics. In this discussion, I will use ‘economism’ to describe any theory or attitude that attributes decisive or conclusive importance to economic considerations in making policy decisions⁷. It is the view that our policy decisions should be ultimately based upon their expected economic consequences. And it is, in the present context, the view that we should not value freedom as an end in itself, but primarily as a means to prosperity.

Economism, thus understood, is not a theory in economics, but the *philosophical* stance that economic facts, interests, and goals are the facts, values, and goals that matter most. This stance, however, is often bolstered by the claim that the study of economics is a *science*, and that its theories and predictions have the cognitive authority that only a science can have. The most obvious proponents of economism are economic reductionists who believe that all facts, values, interests, and goals can ultimately be defined in economic terms — or, in other words, that economic facts, values, interests, and goals are the *only* ones that really exist.

Marx is probably the best known and most influential proponent of this view, and the prevalence of economism in contemporary thought is undoubtedly due to his influence. But if I am right, then Hayek retained elements in his own thinking that are economic as well.

Thus, Hayek argued that scientific study has shown that socialist economic programs and aims are both empirically and logically mistaken, and he said that the fact that the socialists were wrong about the economic *facts* was crucial to his critique of socialism. Hayek also argued that scientific study has shown that socialist programs — and, in particular, central planning— *cannot* succeed in achieving their aims. But what is, perhaps, more to the point is that Hayek held that freedom is important first and foremost for its economic consequences. Whereas Popper thought that it was wrong to base the rejection of tyranny on economic arguments, Hayek was apparently ‘prepared’ to sacrifice individual freedom, if his analysis of the economic consequences of socialism proved wrong.

⁷ This definition is taken from *Webster's third new international dictionary*. It is interesting that no entry for ‘economism’ appears in *The Oxford English dictionary*.

Freedom or efficiency?

Popper and Hayek were both attracted to socialism as young men, and both, late in life, stated the conditions under which they would accept it. Hayek thus writes, in the opening pages of his last book, *The Fatal conceit*, that:

I am prepared to admit that if socialist analyses of the operation of the existing economic order, and of possible alternatives, were factually correct, we might be obliged to ensure that the distribution of incomes conform to certain moral principles, and that this distribution might be possible only by giving a central authority the power to direct the use of available resources, and might presuppose the abolition of individual ownership of means of production. If it were for instance true that central direction of the means of production could effect a collective product of at least the same magnitude as that which we now produce, it would indeed prove a grave moral problem how this could be done justly. This, however, is not the position in which we find ourselves. (Hayek, *The fatal* 6-7)

And Popper, in an often delivered lecture called “Open society and the democratic state”, expressed the opposite view as follows:

I believe that a free market-economy is more efficient than a centrally planned economy. Yet I hold that it is wrong to base the rejection of tyranny on economic arguments. Even if it were true that a centrally planned state economy is superior to that of the free market, I should oppose the centrally planned economy. I should oppose it because of the likelihood that it would increase the power of the state to the point of tyranny. It is not the inefficiency of communism, against which we should fight, but its inhumanity and its inherent hostility to liberty. We should not sell our freedom for a mess of pottage, or for the promise that we shall obtain the highest possible productivity and efficiency — not even if we could be sure that we can purchase efficiency at the price of liberty. (Popper, *Archives* Box 6)

These counterfactual considerations were not intended —and should not be interpreted— as arguments for socialism. Hayek did not think that socialism could ever be as efficient and productive as the market. And Popper did not think that socialists would resist the temptation to use an increase in state power to curtail freedom. But the difference in their conditions for accepting socialism is striking.

Popper and Hayek both criticized socialism for being less efficient than capitalism and for leading to totalitarianism. But Popper said that he would oppose central planning even if it were more efficient than the market, and Hayek said that he would be prepared to accept it if it could be shown to be at least as productive. Their positions regarding economism, as illustrated in the two passages quoted above, are thus diametrically opposed to each other⁸.

Here, someone might object that Hayek, in the passage that I have quoted, stated the conditions under which he would be willing to accept *central planning*, and not the conditions under which he would be willing to accept *totalitarianism*. But the issue becomes clear once we remember that Hayek thought that centralized planning *logically* entails totalitarianism. And while it is clear that he regarded socialism as both factually and logically impossible, he offered no reason to think that we could—even in this counterfactual situation that he regards as impossible—give the state power to redistribute incomes without simultaneously sacrificing freedom. His argument, on the contrary, would seem to suggest that the totalitarianism that he thinks central planning entails would be acceptable—and that we might actually be *obliged* to set it in motion, despite the “grave moral problem how this could be done justly”—provided only that we “could effect a collective product of at least the same magnitude as that which we now produce”

⁸ I read the passage from Hayek, of course, as saying that we might be obliged to ensure a more equal distribution of incomes if, for example, central planning could be shown to be at least as efficient and productive as the market. Some friends of Hayek, however, have suggested that his reference to ‘certain moral principles’ would rule this interpretation out, since he regarded socialism as immoral. But it seems clear, against this proposed interpretation, that Hayek recognized that some socialists argue for a more equal distribution of income on the basis of what *they*, at least, perceive to be moral principles—and that he criticized them, at least in part, for trying to replace a moral system that had spontaneously evolved through the extended order with one of their own design. He thus writes, in the same context, that: “The demands of socialism are not moral conclusions derived from the traditions that formed the extended order that made civilization possible. Rather, they endeavor to overthrow these traditions by a rationally designed moral system whose appeal depends on the instinctual appeal of its promised consequences” (Hayek, *The fatal conceit* 7).

But quite aside from this, I submit that the passage that we are considering would be unintelligible if we were to assume that it meant anything other than that Hayek was prepared to admit that we might be obliged to adopt socialism *if his economic analysis were mistaken*. For he says, after all, that the distribution of incomes that he has in mind—the one that would conform to ‘certain moral principles’—“might be possible only by giving a central authority the power to direct the use of available resources, and might presuppose the abolition of individual ownership of means of production” (Ibid. 6). But it is difficult to see why an unequal distribution of goods would require central planning and the abolition of private property. And Hayek, in any event, argued that the equal distribution of incomes would require just such measures. There would, finally, be no reason, on the interpretation that I am rejecting, why he should say that “it would indeed prove a grave moral problem how this could be done justly”. But on the interpretation I propose, the reason is obvious. It is that he regards socialism as immoral.

by doing so. It is precisely this suggestion, and the economism upon which it is based, that Popper rejects when he says that “it is wrong to base the rejection of tyranny on economic arguments”, and that “we should not sell our freedom for a mess of pottage, or for the promise that we shall obtain the highest possible productivity and efficiency — not even if we could be sure that we can purchase efficiency at the price of liberty”.

These arguments presuppose different values —the acceptance or rejection of economism— and the fact that Popper and Hayek agreed that planning is both economically inefficient and a cause of totalitarianism is no reason to ignore it. For what is at issue here is the relative value of freedom and efficiency. It is a matter of priority, or what comes first.

The question, in a nutshell, is whether we should value freedom because freedom is valuable, or because it is profitable. It is the question whether we should regard freedom as an end in itself that is valuable for its own sake, or as a means to economic prosperity that we may dispense with if and when it no longer achieves its end.

There is an irony here that should not be ignored.

Popper wrote in his autobiography that he would still be a socialist if socialism could be combined with human freedom (Popper, *Unended quest* 36)⁹. But he did not think that the two were *logically* incompatible, or that tyranny was a *necessary* consequence of planning. Popper also thought that the market is more efficient than a planned economy. But he thought that this was a matter of fact, as opposed to a logical truth, and said that he would continue to oppose a planned economy even if it were more efficient than the market.

Popper argued that planning requires that we invest a great deal of power in the state; that the greater the power, the greater the likelihood of its abuse; and that the likelihood of avoiding tyranny was thus very slim. And he found corroboration for his conclusion in the fact that the socialists he knew regarded planning as both simple and a cure-all — and were thus generally unprepared for its problems.

⁹ Popper told me, toward the end of his life, that he would still call himself a socialist “were it not for the power problem”.

Hayek, on the other hand, thought that planning *entails* totalitarianism. He argued that its redistribution of wealth would necessitate coercive action on the part of the state that conflicts with the very idea of individual freedom. Yet Hayek said that he was ‘prepared’ to accept planning—and, by implication, coercion and the sacrifice of freedom that it entails—if it could only be shown to be at least as productive as the market.

Here, someone might object that my argument places too much weight upon an isolated counterfactual hypothesis, especially since Hayek regarded its antecedent—that planning can be as efficient and productive as the market—as false. My own sense, however, is that Hayek’s economism is not isolated at all, but is actually built into his idea of individual freedom. In what follows, I will first trace Hayek’s ideas of individual freedom and coercion to his economism. I will then argue that Hayek’s idea of freedom suffers from the same ‘economic’ considerations that he brought to bear against planning, and that the ‘economic’ considerations that he brought to bear against planning have since been superseded by developments in technology.

Individual, political, inner, and economic freedom

Hayek wrote that “‘freedom’ refers solely to a relation of men to other men, and the only infringement on it is coercion by men” (Hayek, *The constitution* 12). He thought that we should not, for this reason, confuse freedom with power, or with the alternatives that may enhance our choice of action within a situation. He thus drew a sharp distinction between ‘individual freedom’, which is freedom from coercion and “independence of the arbitrary will of another” (Ibid. 12), and ‘political freedom’, which allows for “the participation of men in the choice of their government, in the process of legislation, and in the control of administration” (Ibid. 13).

Political freedom may be desirable for its own reasons. But Hayek thought that “a free people in this sense is not necessarily a people of free men; nor need one share in this collective freedom to be free as an individual” (Ibid. 13). Individual freedom, by contrast, “presupposes that the individual has some assured private sphere, that there is some set of circumstances in his environment with which others cannot interfere” (Ibid. 13)—so that we are free as individuals to the extent to

which we can follow our own plans and intentions toward ends of our own choice, *as opposed to having to submit to necessities created by others to make us do what they want us to do* (Ibid. 13).

Hayek thought that this kind of freedom, is based upon five basic legal rights: (1) legal status as a protected member of the community; (2) immunity from arbitrary arrest; (3) the right to work at whatever one desires to do; (4) the right to movement according to one's own choice; and (5) the right to own property (Ibid. 20). And he warned that the confusion of individual freedom with freedom as power inevitably leads us to identify it with wealth (Ibid. 17).

Hayek also drew a distinction between individual freedom and what he called 'inner' or 'subjective' or 'metaphysical' freedom. Inner freedom

refers to the extent to which a person is guided in his actions by his own considered will, by his reason or lasting conviction, rather than by momentary impulse or circumstance. (Ibid. 15)

It is closely related to individual freedom, but it depends upon a person's strength of will and not upon whether or not he is coerced by others — though Hayek acknowledged that "they are clearly not without connection", since:

The same conditions which to some constitute coercion will be to others merely ordinary difficulties which have to be overcome, depending on the strength of will of the people involved. (Hayek, *The constitution* 15)

'Economic freedom' is another matter. Hayek did not like the term because he thought that it had been co-opted by the socialists for their own illiberal purposes.¹⁰ He wrote that "political freedom is meaningless without economic freedom" and that economic freedom is "the prerequisite for any other freedom" (Hayek, *The road* 100). But he used the term in an ironic sense, which he underscored by adding that the economic freedom to which he referred was utterly different

¹⁰ This, incidentally, also holds true of 'political freedom'. In *The road to serfdom* Hayek writes that, "to the great apostles of political freedom the word meant freedom from coercion, freedom from the arbitrary power of other men, release from the ties which left the individual no choice but obedience to the orders of a superior to whom he was attached" (Hayek, *The road* 25). And this, of course, seems more closely aligned with what he called 'individual freedom' in *The constitution of liberty*.

from the socialists' use of the term. Freedom in the socialists' sense — the “economic freedom” without which the political freedom already gained was ‘not worth having’” (Hayek, *The road* 25)— is “merely another name for power or wealth” (Ibid. 26). The economic freedom that Hayek said is necessary for any other freedom is just the individual freedom that he associated with the five basic rights mentioned above.

Hayek argued that central planning destroys individual freedom, not political freedom — and that the destruction of individual freedom, not political freedom, is what results in totalitarianism. Political freedom —or democracy— may be meaningless without economic freedom, but it is possible nonetheless. And Hayek not only thought that democracy's ‘political freedom’¹¹ is compatible with totalitarianism, he argued that ‘totalitarian democracy’ is what ‘social’ democracy —as opposed to ‘liberal’ democracy— actually entails (Hayek, *The constitution* 55-6).

Popper agreed that democracy cannot guarantee freedom, and that it is more likely to devolve into tyranny in states that do not already have strong traditions of liberty (Popper, *Archives*). But he also thought, contrary to Hayek, that democracy is “the basis of everything else”. For,

without democratic control, there can be no earthly reason why any government should not use its political and economic power for purposes very different from the protection of the freedom of its citizens. (Popper, *The open society* 127)

The case for individual freedom

Hayek wrote that:

The case for individual freedom rests chiefly on the recognition of the inevitable ignorance of us all concerning a great many of the factors on which the achievement of our ends and welfare depends. (Hayek, *The constitution* 29)

¹¹ I have written ‘political freedom’ here in quotes, because Hayek did not regard it, in and of itself, as real freedom at all. Hayek thought that political freedom is real freedom only when accompanied by individual freedom, which insures the right to own property. This, I think, is a consequence of his idea that economic freedom is a prerequisite for any other freedom.

“What is important is not what freedom I personally would like to exercise but what freedom some person may need in order to do things beneficial to society” (Hayek, *The constitution* 32). And “if the result of individual liberty did not demonstrate that some manners of living are more successful than others, much of the case for it would vanish” (Ibid. 85).

Individual freedom, in other words, is valuable for its economic consequences —because it makes the extended spontaneous order of the market possible, and because the market, by utilizing knowledge that is dispersed among individuals, enables us to solve the so-called ‘calculation problem’, which thus makes possible the economic prosperity that, in turn, enables the survival of large numbers of human beings.

This is why planning conflicts with freedom. Hayek recognized that economic knowledge is dispersed amongst the members of a society; that it is simply too complex for any single mind or central committee of minds to comprehend; and that the market is first and foremost an *epistemological* tool that enables buyers and sellers to signal to each other, via the relative and changing prices of the goods that are offered for sale, the facts about supply and demand. He then argued that the socialists’ attempt to run an economy without a market was doomed to fail, since it would unwittingly deprive them of the very information that they would need for successful planning.

The rule of law

Earlier I said that Hayek’s idea of freedom is based upon legal rights. We are free, according to Hayek, because we are bound by laws that enable us to choose our actions with an understanding of their legal consequences — and, thus, with an assurance that they will not incur unexpected coercive responses from the state. So the rule of law, far from being an impediment to freedom, is actually its necessary condition. This emphasis upon the rule of law may, at first glance, seem entirely independent of economic considerations. It may also seem entirely compatible with the aims of socialism. Hayek, however, thought that the rule of law provides the criterion for determining whether or not an economic measure undertaken by government is compatible with a free system, and that it is actually incompatible with certain economic

aims of socialism, most notably distributive justice. Getting clear as to why he thought these things will give us yet another window on his economism.

Hayek wrote that the rule of law “means that government must never coerce an individual except in the enforcement of a known rule”. The rule of law thus “constitutes a limitation on the powers of all government, including the powers of the legislature” (Hayek, *The constitution* 205) and Hayek thought that many of the laws that have been written by legislatures are actually in violation of it. True laws, for Hayek, must be ‘known and certain’ (Ibid. 208) and they must apply equally to all people who are subject to them (Ibid. 209). He thus argued that government action in a free society must restrict itself to the enforcement of general laws, as opposed to specific commands, which do not apply to everyone and are designed to benefit particular individuals and groups.

Still, it may not be immediately clear why, or how, a proposal to legislate distributive justice—say by a law that “all legal residents of the United States over the age of 21 will be guaranteed a yearly income equivalent to \$50,000; in real 2001 U.S. dollars” — violates the rule of law. The proposed ‘law’ seems to be completely general in form and would thus apply equally to all legal residents of the United States. It does not seem to have the character of a specific command. Nor does it seem to benefit particular individuals or groups as opposed to others. It could, if enacted by the appropriate legislative authorities, be made just as ‘known and certain’ as any other law. Hayek, however, thinks that the rule of law requires more than mere legality, or conformity to the law; and more than even constitutionalism, or the requirement that laws passed by a state are not in violation of its written constitution. The rule of law, over and above this, “requires that all laws conform to certain principles” (Hayek, *The constitution* 205).

It is important, for this reason, to distinguish between laws and the rule of law. The rule of law “is not a rule of the law but a rule concerning what the law ought to be, a meta-legal doctrine or a political ideal” (Ibid. 206). It is not an actual law itself, but a meta-legal principle that provides a normative guide for the conception of proper laws.

This meta-legal principle is the principle of liberalism, interpreted by Hayek to mean

a policy which deliberately adopts competition, the market, and prices as its ordering principle and uses the legal framework enforced by the state to make competition as effective and beneficial as possible. (Hayek, *Individualism* 110)

Hayek thus wrote that “it is of the essence of the rule of law that the private citizen and his property should not in this sense be means at the disposal of government” (Hayek, *The constitution* 214).

Here, it should be clear why Hayek thought that any legislation that would guarantee a set fixed income to all individuals in a society would violate the rule of law. For Hayek held that there can be no hope of guaranteeing fixed incomes unless we replace the market ordering principle with central planning — and that this, by its very nature, would be a violation of the principle of liberalism. It should, moreover, also be clear that Hayek’s appeal to the principle of liberalism in characterizing the rule of law —and, thus, in characterizing individual freedom— is not only fully consistent with economism but is actually driven by it in at least two different ways.

First, the purpose of the law is to carve out an ‘assured private sphere’ that makes free choice of action possible. This is what laws are supposed to do, and it is the reason why they must be general and well-known. Commands given to individuals and groups to perform specific actions may be arbitrary and unexpected. They do not apply to everyone and the consequences for violating them may be unexpected and arbitrary as well. They thus fail to provide a legal framework in which individuals can deliberate and act with the assurance that their actions will not incur unexpected coercive responses from the state. This legal framework is a large part of what Hayek means by an ‘assured private sphere’. But Hayek also regarded private property as an essential prerequisite for such a private sphere. For without the recognition of private property, the individual will not be able to carry out his own plans of action, and hence will not be free.

The second way in which the rule of law is driven by economic considerations becomes clear when we consider its role in placing limitations upon the powers of government, including the legislature.

Hayek thought that the laws in a free society must be written in conformity with the rule of law. This, as we have just seen, means that

they must be general and well-known. It also means that the state must never coerce an individual except to enforce a true law. ‘Government’, Hayek wrote, “must not use its power of coercion to reserve for itself activities which have nothing to do with the enforcement of the general rules of law” (Hayek, *The constitution* 223).

But Hayek thought that legislatures can and do enact all sorts of ‘policy’ rules—which may be both general and well-known—that do not qualify as ‘true laws’. And his judgments as to whether a particular piece of legislature constitutes a true law, as opposed to a rule of policy, typically turn upon whether or not he thinks that the enforcement of the measure would constitute a legitimate use of state coercion. The circularity of reasoning here may be unavoidable. But the point to be made is that Hayek typically uses the economic consequences that a given measure has for the market as a touchstone for determining whether or not it is truly coercive.

Coercion

Hayek defined ‘individual freedom’ as absence of arbitrary coercion, and he said that his definition of ‘freedom’ would ultimately depend upon his definition of ‘coercion’. So the question ‘What is coercion?’ is the heart of the problem. Hayek wrote that:

By “coercion” we mean such control of the environment or circumstances of a person by another that, in order to avoid greater evil, he is forced to act not according to a coherent plan of his own but to serve the ends of another. (Hayek, *The constitution* 20-1)

Coercion is thus something that one human being does to another. This seems intuitive enough. But further analysis of Hayek’s definition shows that it is both too broad and too narrow, and that coercion is far too subjective an idea to bear the weight that Hayek has placed upon it.

Here, the fact that coercion is something that one human being does to another is important. For it occurs, according to Hayek: only if our options are controlled *by other human beings*. So if our options are not controlled by other human beings—if we are merely playing the hand that God has dealt—then we are not coerced. This means that whether or not a person is coerced will depend upon whether or not his circumstances

are controlled by another person. And *this* means that there may well be situations in which we are forced, *but not coerced*, to do something that we do not want to do in order to avoid what we perceive to be a greater evil. But this is not all. For coercion occurs, according to Hayek only when “one man’s actions are made to serve another man’s will, not for his own but for the other’s purpose” (Hayek, *The constitution* 133). And here, the “not for his own but for the other’s purpose” is equally important. Without this clause every act of employment would be an act of coercion. But if my actions are made to serve your will both for your purposes and for my own—if I do your will in order to put bread on my table— then my actions are not coerced. Finally, the fact that it is *actions* that are coerced is crucial. For action, according to Hayek, implies choice, and we cannot be coerced unless we can choose. Indeed, one might say that it is the *choice* between alternative actions that is or is not coerced.

I can still choose how to act even if someone has a gun to my head. But if my behavior is physically forced—if, for example, you physically force my finger to pull the trigger on a gun— then I have not been coerced in Hayek’s sense. For I have no choice and, thus, have not acted if my behavior is physically forced. And if I have not acted, then my actions have not been coerced. It is, however, easy to see how this definition can lead to problems. Many people, contrary to Hayek, would regard cases in which people physically force others to do things against their will as paradigmatic examples of coercion. And it certainly seems strange, when we combine it with Hayek’s claim that the *only* infringement upon freedom is coercion by men (Hayek, *The constitution* 12), to say that we are not coerced in such cases. For surely you infringe upon my freedom if you physically force my finger to move in ways in which I do not want it to move.

Considerations such as these suggest that Hayek’s definition of ‘coercion’ is too narrow. But other considerations suggest that it is also too broad. We coerce people when we force them to act in such a way as to do our will. But if a person is really to *act*, then he must also be able to choose. And if he is able to choose, then he must have options. But if a person has options, then he can always choose *not* to serve our will. And if he can choose not to serve our will—as he must be able to do, on Hayek’s account, if he is really coerced— then what, exactly, does it mean to say that he has been *forced*, let alone that *we* have forced him, to serve our ends *instead* of his own?

Here, Hayek's idea that "coercion occurs when one man's actions are made to serve another man's will, not, for his own but for the other's purpose" is especially problematic. For even if I choose to do your will only in order to avoid a more painful consequence, am I not still serving my own purpose as well? Hayek says that coercion implies,

that I still choose but that my mind is made someone else's tool 'because the alternatives before me have been so manipulated that the conduct that the coercer wants me to choose becomes for me the least painful one'. (Hayek, *The constitution* 133)

This is no doubt true. But it seems to open up a Pandora's Box of epistemological problems about how to determine whether or not a person's actions, beliefs, and choices have actually been coerced. Hayek says that "coercion implies the threat of inflicting harm and the intention thereby to bring about certain conduct" (Ibid. 134).

But he also says that "whether or not attempts to coerce a particular person will be successful depends in a large measure on that person's inner strength", that "the threat of assassination may have less power to turn one man from his aim than the threat of some minor inconvenience in the case of another" (Ibid. 138), and that a weak or very sensitive person may be coerced by a disapproving glance to do what he would not otherwise do. Hayek, in discussing these examples, says that we are concerned with coercion that is likely to affect the normal, average person. But the relativity of coercion, once it is introduced, would seem to render coercion an entirely subjective concept.

Hayek, of course, recognized the problem. But he said that it "does not mean that we ought not to try to prevent all the more severe forms of coercion, or that we ought not to define liberty as the absence of such coercion" (Ibid. 139). He tried, instead, to objectify the notion by talking about the 'normal, average person', and 'degrees of coercion', and its 'more severe forms'. But the subjectivity of coercion nonetheless infects Hayek's entire discussion of freedom. And it is easy to see why. For judgments about which forms of coercion are more severe and which forms are likely to affect the 'normal, average' person will most likely be as subjective as the concept of coercion itself. Using these judgments as touchstones to determine whether or to what extent another individual has been coerced—and, hence, whether or to what extent he

is free— would pose problems even if they were rigorously and entirely consistent. But these problems will only be magnified if the judgments seem arbitrary and inconsistent. And this, unfortunately, is often the case with Hayek. The higher our standards for what counts as coercion, the lower they are for what it takes to be free. And it sometimes seems as if Hayek uses a sliding scale according to which a given action can be coercive or non-coercive depending upon whether or not he approves of it.

Hayek thus argues, on the one hand, that:

The individual provider of employment cannot normally exercise coercion, any more than can the supplier of a particular commodity or service. So long as he can remove only one opportunity among many to earn a living, so long as he can do no more than cease to pay certain people who cannot hope to earn as much elsewhere as they had done under him, he cannot coerce, though he may cause pain. (Hayek, *The constitution* 136)

But he also argues, on the other hand, that labor unions are able to coerce the providers of employment, even though “the coercion which unions have been permitted to exercise contrary to all principles of freedom under the law is primarily the coercion of fellow workers” (Ibid. 269).

So the fact that a worker may seek employment elsewhere means that a provider of employment cannot coerce his employees to accept his conditions—despite the facts that the employer may cause pain and his employee may not be able to find comparable conditions anywhere else. But a labor union can coerce both its members and their employer—despite the fact that the workers can seek employment, and the employers other workers, somewhere else.

The subjectivity of coercion is just the tip of the iceberg. Another, and perhaps more important, problem is the relationship between freedom and power.

Freedom, power, and 'true coercion'

Hayek, as we have already seen, drew a sharp distinction between freedom and power. He said that confusing the two "inevitably leads to the identification of liberty with wealth" (Hayek, *The constitution* 17), and then to demands for its redistribution - or 'social justice'. He thought that freedom and power are two desirable things, but that they are also two entirely different things:

Whether or not I am my own master and can follow my own choice and whether the possibilities from which I must choose are many or few are two entirely different questions. The courtier living in the lap of luxury but at the beck and call of his prince may be much less free than a poor peasant or artisan, less able to live his own life and to choose his own opportunities for usefulness. Similarly, the general in charge of an army or the director of a large construction project may wield enormous powers which in some respects may be quite uncontrollable, and yet may well be less free, more liable to have to change all his intentions and plans at a word from a superior, less able to change his own life or to decide what to him is most important, than the poorest farmer or shepherd. (Ibid. 17)

But it is not at all clear that these examples actually support the distinction between freedom and power. For the restrictions on the freedom of the general result from his having a superior with even greater power than he has. And while the courtier may be living in the lap of luxury, the lap in which he is living is not his own. He has neither independent wealth nor independent power — and he would most probably not be courtier to a prince if he did, let alone be living at his constant beck and call.

The issue becomes more complicated, and more interesting, when Hayek tries to distinguish between power and coercion. Hayek wrote that:

It is not power as such —the capacity to achieve what one wants- that is bad, but only the power to coerce, to force other men to serve one's will by the threat of inflicting harm. There is no evil in the power wielded by the director of some great enterprise in which men have willingly united of their own will and for their own purposes. It is part of the strength of civilized society that, by such voluntary

combination of effort under a unified direction, men can enormously increase their collective power.

It is not power in the sense of an extension of our capacities which corrupts, but the subjection of other human wills to ours, the use of other men against their will for our purposes. (Hayek, *The constitution* 134-5)

Here, Hayek begins by saying that it is not power as such that is bad, but only the power to coerce. But he then says that it is not power but coercion that *corrupts*. And this is an entirely different matter.

Hayek is not saying that *all* uses of coercive power are bad, but that *only* uses of *coercive* power are bad. But the fact remains that wrongful coercion is *itself* a corrupt use of power. I attempt to coerce others when I threaten to inflict harm upon them in order to force them to serve my will. I need not actually have the power to inflict harm for my attempt to work. They need only believe that I have it. But the power to inflict harm no doubt provides me with an opportunity to coerce others. And if I have this power, and use it in this way, then I have already been corrupted by it.

There may be no evil in power *per se*. But to speak of 'the power to coerce' as if it was somehow different from 'power as such' is misleading if it suggests that we need not worry about power as such. It is not that the power to coerce corrupts but power as such does not. And it is not that coercion is a different kind of power. The power to coerce is power pure and simple. It becomes coercive when *we use it* to force other people to serve our will instead of their own. And someone who uses his power in order to coerce others has already been corrupted by it. I do not want to deny that successful attempts at coercion may make one more corrupt than he already is — or to suggest that one must actually be able to inflict harm in order to coerce others. For one may often coerce through bluffs that he cannot fulfill. And a manipulator may, in this way, all too easily progress to a monster. But power as such may corrupt simply by tempting a person to use it to coerce others in order to get his own way.

Popper, contrary to Hayek, knew that *all* power tends to corrupt. This, in fact, was a large part of his critique of Marx's economism. He wrote that Marx and his followers were both inconsistent and mistaken to exaggerate the role of economic power in relation to political and physical power:

Their argument runs: he who has the money has the power; for if necessary, he can buy guns and even gangsters. But this is a roundabout argument. In fact, it contains an admission that the man who has the gun has the power. And if he who has the gun becomes aware of this, then it may not be long before he has the gun and the money. (Popper, *The open society* 127)

Hayek, no doubt, agreed. He wrote that:

True coercion occurs when armed bands of conquerors make the subject people toil for them, when organized gangsters extort a levy for “protection”, when the knower of an evil secret blackmails his victim, and, of course, when the state threatens to inflict punishment and to employ physical force to make us obey its commands. (Hayek, *The constitution* 137)

But Popper did not deny, and did not want to deny, that economic power is dangerous and corrupting. He thought that the poverty of Marx’s economism is that it saw economic power everywhere and regarded it as more fundamental than physical power and the power of the state. But this is because “all power, and political power at least as much as economic power, is dangerous” (Popper, *The open society* 129) and because Marx had underestimated the role that democracy, or what the Marxists denigrated as ‘merely formal freedom’, could play in controlling it. Popper thus argued that an unrestrained capitalism that “develops institutions for the control of guns and gangsters but not of the power of money is liable to come under the influence of this power” (Ibid. 129). And what emerged was a critique of economism that portrays economic power as dangerous, but as *no more* dangerous than either political power or the physical and psychological forces upon which they are ultimately based.

There is more at stake here than the definitions of ‘individual freedom’ and ‘true coercion’. Whether or not a particular action is a threat to our freedom will depend, on Hayek’s account, upon whether or not it is truly coercive. But Hayek thought that the primary function of government is to protect the individual against arbitrary coercion. And if this is true, then what hangs in the balance are the kinds of coercion that its citizens can legitimately expect their government to redress.

True coercion, according to Hayek, will require ‘very exceptional circumstances’ (Hayek, *The Constitution* 135). For,

so long as the services of a particular person are not crucial to my existence or the preservation of what I most value, the conditions he exacts for rendering these services cannot properly be called “coercion”. (Ibid. 136)

And,

unless a monopolist is in a position to withhold an indispensable supply, he cannot exercise coercion, however unpleasant his demands may be for those who rely on his services. (Ibid. 136)

But even a ‘normal, average person’ might regard circumstances less exceptional than these coercive.

Popper criticized the nineteenth century apologists of unrestrained capitalism who appealed to the idea of freedom—and the slogan ‘equal and free competition for all’—to cynically and hypocritically defend the coercion of workers that Marx described as ‘exploitation’:

Marx lived, especially in his younger years, in a period of the most shameless and cruel exploitation. And this shameless exploitation was cynically defended by hypocritical apologists who appealed to the principle of human freedom, to the right of man to determine his own fate, and to enter freely into any contract he considers favourable to his interests.

Using the slogan ‘equal and free competition for all’, the unrestrained capitalism of this period resisted successfully all labour legislation until the year 1833, and its practical execution for many years more. The consequence was a life of desolation and misery which can hardly be imagined in our day. (Popper, *The open society* 122)

I do not want to suggest that Hayek himself was either cynical or hypocritical in his account of freedom. But Popper thought that an employer can intimidate, and exploit, and, in a word, *coerce* his employees even under very normal circumstances. He also thought that we should try to find a *political* solution to the problem:

Even if the state protects its citizens from being bullied by physical violence (as it does, in principle, under the system of

unrestrained capitalism), it may defeat our ends by its failure to protect them from the misuse of economic power. In such a state, the economically strong is still free to bully one who is economically weak, and to rob him of his freedom. Under these circumstances, unlimited economic freedom can be just as self-defeating as unlimited physical freedom, and economic power may be nearly as dangerous as physical violence; for those who possess a surplus of food can force those who are starving into a 'freely' accepted servitude, without using violence. And assuming that the state limits its activities to the suppression of violence (and to the protection of property), a minority which is economically strong may in this way exploit the majority of those who are economically weak.

If this analysis is correct, then the nature of the remedy is clear. It must be a *political* remedy—a remedy similar to the one which we use against physical violence. (Popper, *The open society* 124-5)¹²

Here, the idea that those who possess a surplus of food can force those who are starving into a 'freely' accepted servitude is really the crux of the matter. A Socrates or a Gandhi might have the inner strength to freely choose starvation over servitude. But a 'normal, average person' does not. Such 'freely' accepted servitude, for him, is more typically coercion. And the fact that freedom, if left unconstrained, can lead to servitude was no surprise to Popper. It was, on the contrary, part and parcel of what he called 'the paradox of freedom'.

The paradoxes of freedom

Popper believed that we should try to maximize the freedom of each within the limits imposed by the freedom of all — and that such maximization of freedom actually requires its limitation. This idea, which Hayek shared, was an attempt to solve what he called 'the paradox of freedom'.

¹² Popper, in a footnote to this passage, acknowledged that his analysis would not be correct if there were perfect competition between the buyers of labor on the labor market. But he rejected the assumption of perfect competition as unrealistic.

The paradox of freedom, simply put, is the idea that freedom without restraints enables the strong to bully the weak, and ultimately leads to tyranny as the strong try to rob the weak of their freedom¹³.

Popper traced the paradox of freedom to Plato's belief that "the probable outcome of too much freedom is only too much slavery in the individual and the state" (Plato 564 a.).

Plato wrote that "the climax of popular liberty" is attained "when the purchased slaves, male and female, are no less free than the owners who paid for them" (Ibid. 563 b.). And he used this idea to argue that we should eschew general freedom and democracy, saying that "tyranny develops out of no other constitution than democracy—from the height of liberty, I take it, the fiercest extreme of servitude" (Ibid. 564 a.). Popper agreed that unrestrained freedom leads to its destruction. But the paradox, he thought, shows only why freedom and democracy should not be left unrestrained. The state, he said, "should limit freedom to a certain extent, so that everyone's freedom is protected by law" (Popper, *The open society* 124). The issue between Popper and Hayek is 'To what extent?'

The question 'To what extent?' has at least two dimensions here: whether such restrictions should apply to our *economic* freedoms and, if so, under what conditions. The first seems clear enough. Popper acknowledged that these restrictions to our freedoms were originally meant to apply to the realm of brute force and physical intimidation, but he argued that they should now also be applied to the economic realm and that unrestrained capitalism should give way to *economic interventionism* (Ibid. 125).

Hayek apparently agreed that we should be able to restrict freedom in the economic realm, for he would otherwise be unable to regard a socialist state monopoly in employment, or the pressures that labor

¹³ Open society is associated with freedom, tolerance, and democracy. But Popper thought that each of these ideas, left uncontrolled, could lead to paradox. The paradox of democracy is closely related to the paradox of freedom. It is the possibility that the majority may decide that a tyrant should rule, or that a group of electors may vote to suspend its own privilege to vote. And the two are also related to the paradox of tolerance, which Popper used to illustrate with the story of the tolerant tribe that extended its tolerance even to the den of man-eating tigers that lived close by. The tigers were happy to be tolerated, but the tribe and its tolerance slowly disappeared. The idea behind these paradoxes is that an excess of freedom, tolerance, or democracy can destroy, respectively, freedom, tolerance, and democracy.

unions exert upon workers, as unacceptably coercive¹⁴. But they seem to disagree when it comes to the conditions under which such restrictions should be made.

Popper argued that the state should restrict the freedom of employers for the simple reason that employers need not resort to actual threats of physical violence in order to coerce those who are starving to accept their conditions. But Hayek argued that an individual employer *cannot* normally coerce workers to accept his conditions so long as he is but one employer among many. Still, one might think that this does not quite answer the question ‘To what extent?’ For what does Hayek mean by ‘normally’? And what does he mean by ‘many’?

Popper, as we have seen, acknowledged that we might not need economic interventionism if there were perfect competition among the buyers of labor in the labor market. And Hayek argued that a monopoly in this market would produce “unlimited powers of coercion”, quoting Trotsky’s observation that “in a country where the sole employer is the state, opposition means death by slow starvation” (Hayek, *The constitution* 137).

So one might want to ask how many potential employers must there be for a worker to be free from such coercion, and whether or not Hayek would regard the choice that Popper poses between work and starvation as normal.

But there need be no doubt about Hayek’s position here, or that it stands in stark contrast to Popper’s. For Hayek, in arguing that withholding a benefit is coercive only when there is a monopoly of an essential service, wrote that:

Even if the threat of starvation to me and perhaps to my family impels me to accept a distasteful job at a very low wage, even if I am “at the mercy” of the only man willing to employ me, I am not coerced by him or anybody else. (Ibid. 139)

¹⁴ I do not, of course, want to equate these two things. I think that a socialist state monopoly on employment can be coercive precisely because it is a state monopoly. Labor unions can also be coercive, but they are significantly different and I do not see why *Hayek* should—or, indeed, how he *can*—regard them as coercive. For even when they are able to achieve a monopoly within a certain field, a worker can always ‘freely’ choose not to work in that field.

This passage—which occurs in the paragraph that immediately follows his quotation of Trotsky—, is stunning. One may, of course, read it as an encomium to freedom—intended, perhaps, to bolster the flagging spirits and ‘inner strength’ of those who feel trapped inside miserable ‘dead-end’ jobs from which they cannot escape. And reading it this way might even inspire some people to try to change their lives.

But it is also possible to see a somewhat different paradox of freedom in it—a paradox pertaining not to the consequences of unlimited freedom, but to Hayek’s very different analyses of what many would regard as somewhat similar cases. For it is difficult, when all is said and done, to understand why we should not regard normal, average people who find themselves in such situations as coerced. And it is also difficult to understand how Hayek can reconcile his claim that they are not with his idea that workers in a socialist state are coerced because opposition to the state means starvation.

Freedom, calculation, and the law

Earlier I said that Hayek’s notion of individual freedom is based upon the rule of law, and that it is impossible to realize it for the very same reason Hayek thought that socialism cannot succeed. Hayek said that his critique of socialism depended upon the economic facts—and upon the fact that our economic knowledge is dispersed amongst the members of a society and far too complex for any one mind, or central committee of minds, to understand.

Hayek said that the market is first and foremost an epistemological tool, and that socialism, by depriving itself of this tool, had condemned itself to failure. But Hayek did not notice that a very similar problem, if not the very same problem, applies to his notion of individual freedom. For the legal codes of really existing states are dispersed in libraries and far too complex for any single mind or committee of minds to understand. Here, one need only think of the Internal Revenue Code—which is, of course, just one small part of US federal law—to recognize the fact. So even though the laws that bind us may be entirely general and publicly accessible, they do not seem capable of providing us with that ‘assured private sphere’ that Hayek thought was necessary for individual freedom.

For we simply cannot foresee the legal consequences of all of our actions clearly enough to be able to plan our lives in such a way as to avoid the possibility of unexpected state coercion. And if our freedom as individuals really depends upon our ability to plan our lives in this way, then we would have little choice but to conclude that we are not really free.

The recognition that market prices, and the competitive system upon which they are based, are epistemological tools was a brilliant insight¹⁵. But Hayek may have been too quick to think that we would never be able to create better ones. He wrote that competition in the realm of commodities means that “as much will be produced as we know to bring about by any known method” — casually adding that “it will of course not be as much as we might produce if all the knowledge anybody possessed or can acquire were commanded by some one agency, and fed into a computer”, but immediately dismissing the idea, saying,

we do injustice to the achievement of the market if we judge it, as it were, from above, by comparing it with an ideal standard which we have no known way of achieving.
(Hayek, *New studies* 185)

It is stunning to think that Hayek wrote these words in 1968. For the fact of the matter is that we have, since 1968, developed another epistemological tool that is potentially as well-equipped to carry out the signaling function of the market — if not actually better equipped to do so even now.

I am talking, of course, about the personal computer and our electronic information technologies, which have given rise to the internet. What Hayek said was certainly true in 1968, when the information technologies that would be necessary for socialist calculation were still just a utopian dream. But the idea that we might, very soon, while sitting in our own living rooms, type out, on a monthly or weekly or even daily basis, exactly what we need to survive, and in what quantities, and with what priority no longer seems so far-fetched at all. I do not, myself, regard

¹⁵ Hayek attributes the insight to Ludwig von Mises: “The distinction of having first formulated the central problem of socialist economics in such a form as to make it impossible that it should ever again disappear from the discussion belongs to the Austrian economist Professor Ludwig von Mises. In an article on *Economic Calculation in a Socialist Community*, which appeared in the spring of 1920, he demonstrated that the possibility of rational calculation in our present economic system was based on the fact that prices expressed in money provided the essential condition which made such reckoning possible” (Hayek, *Collectivist economic* 33).

this as a reason to embrace socialism. But I do think that it should give market proponents of economism pause for thought.

Hayek thought that individual freedom is valuable for its economic consequences, and that its economic consequences are valuable for human survival. We have, since Darwin, grown accustomed to thinking that valuing something for the sake of survival is as close as one ever gets to valuing it for its own sake. But this was not the reason why those who cried 'Give me liberty or give me death' valued it. And it was not the reason why Socrates chose to drink the hemlock. The appeal of economism is that it gives a clear rationale for freedom, and one to which 'the normal, average person' can relate. But this rationale is cold comfort to anyone who has to choose between freedom and survival. And the poverty of economism is that the value that it places upon freedom ultimately depends upon empirical facts — upon whether the market is more efficient than central planning, and more able to sustain the existence of human beings 'in large numbers' than other economic arrangements — whose truth, like that of all empirical facts, is, Hayek notwithstanding, contingent upon particular circumstances that may change over time.

This, in a nutshell, means that those who value freedom primarily for its economic consequences may come to repudiate it if and when they feel that the economic benefits of freedom are no longer so obvious, or if and when they discover how to acquire those economic benefits in other ways.

Our information technologies may help to solve the calculation problem. But they will not obviously protect us against tyranny. Nor will they help us to carve out an 'assured, private sphere'. This is not simply because our laws are far too complex for any computer to understand. The fact of the matter is that our judicial process relies not so much upon calculation as judgment. Our laws may be written in books, but they must be interpreted by judges and juries. And it is all too well-known that judges and juries may render decisions that the parties to a dispute might never expect. Indeed, my own sense is that our freedom consists less in our ability to calculate the legal consequences of our actions, than in the fact that we can write our own laws—including our laws for writing laws—and rewrite them whenever we see fit. It consists, in other words, in our ability to practice what Popper called 'piecemeal engineering'.

Rationality as piecemeal engineering

Adam Smith taught that actions motivated by self-interest can benefit society, and Hayek that attempts to benefit society can easily backfire. This idea—that the consequences of our actions may be different from, and even opposed to, the intentions that motivated them—may be the single most important contribution that Economics has made to our understanding of the social world. Hayek took it to mean that we should refrain whenever possible at deliberate attempts at economic intervention. Popper also taught that our actions may have unintended and undesired consequences. He thus thought that the socialists' desire to help the poor and to eliminate unemployment was one of the best motives of his time. But he also thought that their *collectivism* is a threat to freedom. Popper described "the problem of the unintended consequences of our actions" as "the fundamental problem of the social scientist" (Popper, *The myth* 128) and said that "*the main task of the theoretical social sciences... is to trace the unintended social repercussions of intentional human actions*" (Popper, *Conjectures* 342)¹⁶. But he disagreed with Hayek about whether and to what extent the fact that our actions have unintended and undesired consequences should lead us to eschew deliberate and conscious attempts at intervention. This is the problem of rationality.

Popper wrote to Hayek, early on, that he agreed "entirely that we cannot 'plan' civilization, and especially the growth of reason" (Popper, *Letter to Hayek*), and that social institutions and 'formations' such as language and the market are 'grown', as opposed to consciously invented. He said that "it would be infinitely better to leave them untouched rather than to tamper with them in a collectivist manner" (Ibid.). But he also said that he saw no reason to regard such institutions as 'sacrosanct' or immune to change.

On the contrary, Popper told Hayek that:

It makes a tremendous difference whether one merely emphasizes that interventionism is bad, or whether one emphasizes that we have only the choice between various forms of interventionism, and advocates one that is based on a conscious liberal and humanitarian policy. (Ibid.)

¹⁶ Popper's italics.

He said that “anything that looks like a general attack on interventionism makes this union impossible, for it is rightly felt, by socialists, as an impossible and undesirable wish to return to *laissez faire*”, and that:

This is why an explicit recognition of the need for some interventionism is not only necessary in the interest of clarity, but also in the interest of that union in the camp of freedom which is necessary if collectivism is to be avoided. (Ibid.)

And he underscored the point with regard to unemployment, saying:

I do not think that mass unemployment is simply due to a clumsy interventionism (even though it might be aggravated by it), and I feel that it is just as necessary to emphasize the need of rational interference with “formations” for the avoidance of unemployment as it is to emphasize the danger of collectivism. In fact, it is my conviction, at present at least, that, if we do not *at the same time* emphasize both, and if we present the matter too much as an alternative between scientific rationalism and a humility which considers these formations as sacrosanct, then *freedom will be lost*. (Popper, *Letter to Hayek*)

Popper thought that Hayek had failed to emphasize the need for intervention, and that the emphasis that he placed upon the fact that socialist ideals often produce the opposite of what they are striving for might too easily be interpreted as a new apology for *laissez faire*, and might, in the end, lead to ‘an even greater tragedy’ – namely:

that our own ideals may also produce the opposite of what we have been striving for, by failing to win the confidence of those misguided idealists who press (together with less idealistic pressure groups) for collectivism. (Ibid.)

Economism and totalitarianism

If what I am saying is true, then Hayek’s rejection of socialism was based upon his economism. It was based, more specifically, upon his idea that considerations such as efficiency and productivity are ultimately decisive for the choice of an economic system, together with his idea that an economic system without a market lacks an indispensable

epistemological tool for determining what goods are needed and desired in a society. But Hayek also argued that socialism entails totalitarianism. He wrote in *Freedom and the economic system* that:

The ultimate decision for and against socialism cannot rest on purely economic grounds, and cannot be based merely on the determination of whether a greater or smaller output of society is likely to be obtained under the alternative systems in question. The aims of socialism as well as the costs of its achievements are mainly in the moral sphere. The conflict is one of ideals other than merely material welfare, and the difficulty is that these conflicting ideals still live together in the breasts of most people without their being aware of the conflict. It is on considerations like those discussed here that we shall have to base our final choice. (15-6)

And he developed this argument, which seems at odds with his later statement that we might be obliged to institute planning if his economic analysis is mistaken, in *The road to serfdom*.

Hayek was no doubt referring to the socialists' and capitalists' competing ideals of freedom when he said that the decision for or against socialism must ultimately be based on the resolution of competing moral ideals. And this might suggest that he was, contrary to what I have been arguing, just as opposed to economism as Popper — were it not for that fact that I have already traced his concept of freedom and his judgments regarding coercion to his underlying economism.

I have, however, yet to explain the role that his argument that socialism entails totalitarianism plays in his work.

Hayek dedicated *The road to serfdom* to 'The Socialists of all Parties'. "It took shape", he tells us, "as a warning to the socialist intelligentsia of England" (Hayek, *The road*) that freedom is in conflict with the socialist means that they proposed to realize it. But Hayek was no longer a socialist when he wrote these words. And I would like to suggest that he based his own rejection of socialism on the calculation argument, but rightfully thought that many socialists would not regard it as decisive because they valued freedom more than economic prosperity and were convinced that capitalism poses a serious if not an insurmountable threat to it.

If I am right, then Hayek offered *Freedom and the economic system* and *The road to serfdom*—both of which point to the problem of calculating economic need without a market while attempting to show that central planning, contrary to the aims and ideals of socialism, unwittingly leads to totalitarianism—as arguments to persuade those who would still be willing to support socialism, even if it is less efficient and productive than the market, because they thought that it would enhance their freedom.

But what about the claim, that socialism entails totalitarianism?

The road to serfdom did little to persuade the socialists to whom it was addressed¹⁷. It is difficult, moreover, to see how socialists who valued freedom could possibly have been persuaded by it—any more than capitalists who valued freedom could have been persuaded by Marx's idea that freedom demands the destruction of the market.

For when all is said and done, Hayek's version of economism, like its Marxist cousin, reduces the concept of freedom to an underlying economic reality. Their ideals of freedom are different. But Marx and Hayek both argue that freedom logically requires the right economics.

I think that Hayek was right to say that the socialists used 'economic freedom' as another name for economic power, and to try drawing a distinction between the two. But I also think that Marx was free to use 'economic freedom' to refer to economic power—and that socialists have as much right to use it to refer to the equal distribution of wealth as we have to use it to refer to the market. These arguments ultimately turn upon alleged facts about the meanings of words. They thus depend upon a linguistic approach to philosophy that Popper criticized throughout his life and did his best in his own work to avoid.

Hayek, by adopting this approach, could argue for a stronger conclusion than Popper would ever have proposed. He could say that planning not only poses a threat to freedom, but that the two are *logically* incompatible and cannot possibly co-exist.

¹⁷ Hayek wrote to Popper on 10 July 1944, that its success is not yet among the liberals but "almost exclusively among the conservatives" (*Hayek to Popper*, 10 July 1944). And he tells us, in a foreword added to a 50's edition of the book, that: "...in America the kind of people to whom this book was mainly addressed seem to have rejected it out of hand as a malicious and disingenuous attack on their finest ideals; they appear never to have paused to examine its argument. The language used and the emotion shown in some of the more adverse criticism the book received was indeed rather extraordinary. But scarcely less surprising to me was the enthusiastic welcome accorded to the book by many whom I never expected to read a volume of this type—and from many more of whom I still doubt whether in fact they ever read it" (Hayek, *The road*).

He could, in this way, argue that freedom is actually *precluded* by planning and *predicated* upon an extended order of economic cooperation that is for the most part spontaneous and undesigned. But if the conclusion he drew was stronger, the argument he used was not. For it is convincing, in the end, only to those who are ready to accept his definitions of these terms.

Hayek's arguments work only by insisting that we understand 'freedom' in a certain way — where the truth is that each of us will typically, and rightly, understand it in the way that best protects his own freedom. Marxists no doubt equated freedom with economic power. But Hayek, if I am right, does essentially the same thing. Some people want economic power to buy bread and butter. Others want it to buy shares in Microsoft. But I do not think that it clarifies anything when we pretend that 'freedom' means freedom when we use it to purchase one of these things, and power when we use it to purchase the other. And regardless of whether we are Marxists or free marketeers, when we equate freedom with economic power, we lose sight of what our economic power was supposed to protect in the first place.

The poverty of economism revisited

Proudhon was content to live in poverty, so long as it did not restrict his freedom. He proclaimed himself an 'anarchist', by which he meant an opponent of authoritarian government. And he is famous for having said that 'Property is theft', by which he meant to criticize exploitation. But Proudhon also criticized communism, saying that it, as much as 'unreformed property', would destroy human freedom by depriving an individual of the right to control what he had made.

Marx praised Proudhon's i.e., *What is Property?* calling it 'the first decisive, vigorous, and scientific examination' of the institution. But he wrote his *Poverty of philosophy* in bitter criticism of Proudhon's *Philosophy of poverty*. And his criticism resulted in an equally bitter split between the so-called 'libertarian' and 'authoritarian' versions of socialism.

There can really be little doubt that Marx was aiming at freedom. But there is no doubt at all that he missed his goal. Marx criticized capitalism for its exploitation of its workers. But he and Lenin thought that freedom would somehow depend upon centralization, and upon a 'dictatorship

of the proletariat'. It is easy enough to see a problem in the idea that freedom depends upon dictatorship. Trotsky, we are told, predicted it early on. He is reported to have said that "these methods lead, as we shall yet see, to this: the party organization is substituted for the party, the Central Committee is substituted for the party organization, and finally the 'dictator' is substituted for the Central Committee".

But 'the majority' of Marxists somehow believed that a dictatorship would be acceptable so long as it was a dictatorship of the workers, by the workers, and for the workers. And so it is easy today to forget that there ever was such a thing as a libertarian version of socialism, and that even Marxists thought of themselves as fighting for freedom. Economic power, however, will always be economic power. And there is now a well-known joke that is told by 'capitalists' and Marxists alike. "Capitalism", it begins, "is a corrupt economic system in which man is exploited by his fellow man". "And with communism", it continues, "It is exactly the other way around".

Some proponents of economism tried to run a state without a market. Others would like to run a market without a state. But it is also easy enough to see the problems with unregulated capitalism. The fundamental problem is its tendency toward monopoly—or a 'dictatorship of the bourgeoisie'—which can restrict our freedom just as much as planning by the state. But there is also fraud, collusion, insider trading, and simple theft. Indeed, the 'free market', as it really exists, has as much to do with the invisible pocket as it does with the invisible hand. We can always say that this is a perversion of the free market. But the truth is that the free market is a perversion of reality. And just as the Soviets had to face up to their 'really existing socialism', the West will sooner or later have to face up to its 'really existing markets'. For it is fairly clear, despite all that talk about *laissez-faire*, that the 'freedom' of these markets depends largely upon their regulation by the state.

And this, too, is the poverty of economism. It begins by telling us that freedom depends upon economics, so that we must first have the right economics, and the right economic science, in order to consider ourselves free. It then chastises us, when the 'right economics' fails to produce freedom, for not having implemented it in its purified form. It reminds us, when even the purified form doesn't work, that having the right economics is a necessary but not a sufficient condition for freedom. And it then chides us, if pushed a just a little bit further, for not understanding that it's all just an idealization, and perhaps not even necessary at all.

My own sense is that power —economic or otherwise— is not so much a component of freedom as a way of protecting it. A free society may want to protect itself with a strong military. And this is one of the reasons why we want a strong economy. But the power that protects our freedom should not be confused with our freedom itself. There is no such thing as a purely defensive weapon. And any power that is capable of protecting our freedom is also capable of destroying it. The truth is that all power tends toward totalitarianism and is thus a danger to our freedom — and that we see this danger most clearly when the power that someone else refers to as his ‘freedom’ is used to prevent us from achieving our own economic and non-economic ends.

Popper argued that planning requires the accumulation of power on the part of the state, and that power thus accumulated could all too easily be abused. But he did not think that such abuse was *necessary*. He argued, instead, that it was *more likely* to occur under socialism, because socialists have a kind of uncritical self-confidence and belief in their own infallibility, and that this infallibility makes it difficult for them to acknowledge their mistakes. He thought that the socialists naïve confidence would lead them to ignore indications that their economic programs were not working — and, in the end, to impose them by force ‘for their own good’ upon people who did not really want them and did not clearly benefit from them.

These are the reasons why he opposed economism in the 30’s and 40’s. And they are, I suggest, good reasons why we should oppose it today.

BIBLIOGRAPHICAL REFERENCES

Hayek, Friedrich A. “Freedom and the economic system”. *Public Policy Pamphlet*. Chicago: University of Chicago Press, 1939. Print.

---. *The road to serfdom*. Chicago: University of Chicago Press, 1944. Print.

---. *Individualism and economic order*. Chicago: University of Chicago Press, 1948. Print.

---. *Collectivist economic planning: Critical studies on the possibilities of socialism*. London: Routledge & Kegan Paul, 1975. Print.

---. “Whither democracy”. *New studies in philosophy, politics, economics and the history of ideas*. London: Routledge & Kegan Paul, 1978. Print.

---. *The constitution of liberty*. Chicago: University of Chicago Press, 1978. Print.

---. *New studies in philosophy, politics, economics and the history of ideas*. Chicago: University of Chicago Press, 1985. Print.

---. *The fatal conceit: The errors of socialism*. Chicago: University of Chicago Press, 1991. Print.

---. *Hayek on Hayek: An autobiographical dialogue*. Ed. Stephen Kresge and Leif Wenar. Chicago: University of Chicago Press, 1994. Print.

Notturmo, Mark A. *Science and the open society: The future of Karl Popper's philosophy*. Budapest: Central European University Press, 2000. Print.

Popper, Karl R. *The open society and its enemies*. London: Routledge & Kegan Paul, 1945. Print.

---. "Prediction and prophecy in the social sciences". *Conjectures and refutations: the growth of scientific knowledge*. London: Routledge & Kegan Paul, 1963. Print.

---. *Unended quest: An intellectual autobiography*. London: Routledge & Kegan Paul, 1992. Print.

---. *The myth of the framework: In defense of science and rationality*. Ed. Mark Amadeus Notturmo. London: Routledge & Kegan Paul, 1994. Print.

---. "The open society and the democratic state". *The Karl Popper Archives*, Box 6, File 6.

REASONS AND TWO KINDS OF FACT*

RAZONES Y DOS TIPOS DE HECHO

TONI RÖNNOW-RASMUSSEN

Lund University, Suecia. toni.ronnorasmussen@fil.lu.se

RECIBIDO EL 10 DE MARZO DE 2012 Y APROBADO EL 24 DE MAYO DE 2012

RESUMEN ABSTRACT

Las razones son hechos, es decir, están constituidas por hechos. Este planteamiento de "razón *qua* hecho", tiene mucho respaldo en la literatura reciente. El presente artículo aborda algunas cuestiones que surgen cuando aplicamos esta idea a la distinción entre razones normativas neutrales al agente y razones normativas relativas al agente. Principalmente consideraré dos puntos de vista sobre la naturaleza de los hechos. Según un punto de vista popular, que concibe los hechos como entidades abstractas, la dicotomía neutral-relativa no suele considerarse necesariamente problemática.

Así, en el llamado *enfoque fino*, es posible que algunas razones *qua* hechos sean neutrales al agente y algunas sean relativas al agente. En un segundo enfoque, menos popular, el llamado *enfoque grueso*, los enunciados de razón se refieren a entidades concretas, i.e. son entidades gruesas más que aspectos finos de tipo proposicional las que constituyen razones.

En el desarrollo del artículo, una vez que examinemos algunas de las aparentes ventajas del enfoque fino, argumentaré que resultan problemáticas. Pero, más importante aún, a pesar de algunas ventajas más o menos evidentes, el enfoque fino es inferior al enfoque grueso, por lo menos en algunos aspectos importantes. En primer lugar, el enfoque fino, trivializa el debate sustancial sobre si existen razones neutrales al agente o razones relativas al agente. En segundo lugar, se argumenta que el enfoque fino implica que el agente enfrenta un sinfín de razones; si no se puede resistir esta implicación, el enfoque delgado resulta estéticamente menos atractivo que el enfoque grueso. Sin embargo, el enfoque grueso no es perfecto y al final discutiré un par de objeciones al mismo.

PALABRAS CLAVE KEY WORDS

Razones neutrales al agente, razones relativas al agente, razones completas, razones normativas.

Reasons are facts, i.e., they are constituted by facts. This "reason *quo* fact" claim is much endorsed in recent literature. This paper addresses some issues that arise when we apply this idea to the distinction between agent-neutral and agent-relative normative reasons. I shall mainly consider two views of the nature of facts. Given a popular view, which conceives of facts as abstract entities, the neutral-relative dichotomy is often regarded as not being particularly problematic.

Thus, on this so called *thin approach*, it is possible that some reasons *qua* facts are agent-neutral and some are agent-relative. On a second, less popular approach, the so-called *thick approach*, reason statements refer to concrete entities; i.e., it is thick entities rather than thin propositional-like features that are constitutive reasons.

In the course of the paper, once we examine some of the apparent advantages of the thin approach, I shall argue that they come out as problematic. But, more importantly, despite some more or less obvious advantages, the thin account is inferior to the thicker account in at least some important aspects. First, the thin approach trivializes the substantial debate on whether there are any agent-neutral or agent-relative reasons. Second, it is argued that the thin account implies that agents face endless reasons; if this implication cannot be resisted the thin view appears aesthetically less appealing than the thick account. The thick account is not flawless, though. At the end, I will discuss a couple of objections to it.

Agent-neutral reason, agent-relative reason, complete reason, normative reason.

* I am grateful to: David Alm, Johan Brännmark, Fritz-Anton Fritzon, Wlodek Rabinowicz and Caj Strandberg, for their useful comments. I would also like to thank Lars Bergström, whose comments (on another paper of mine) forced me to think more about the issues discussed in this work. This is a revised version of my paper of the same title, which was published in Frans Svensson and Rysiek Sliwinski (ed.)

Introduction

Reasons are facts, i.e., they are constituted by facts. This “reason qua fact” claim is much endorsed in recent literature. In what follows I would like to address a couple of issues that arise when we apply this idea to practical normative reasons. My interest in this matter dates back to something I recently argued for, namely that we should be cautious about endorsing the distinction between agent-neutral and agent-relative reasons, if we understand this dichotomy in terms of facts (rather than, say, statements or principles). I shall not here repeat my rather extensive argument.

However, I shall discuss some issues that arise for the distinction depending on what sort of view we take on the nature of facts. So although I shall not try to answer the question —‘can we make sense of the dichotomy if we have normative practical reasons qua facts in mind?’- my contribution will at least outline some issues that need to be addressed before we answer this question.

En passant, the questions that I will address are of a magnitude that could easily discourage even the bravest of inquirers. For this reason, I would like to make it clear at the outset that since, inevitably, I will be cutting corners I have no illusions of settling anything at all when it comes to the issues that I will discuss.

Without going into too much detail, we can distinguish in broad terms between two (or perhaps even three) important views on the nature of facts. Given in particular a view that conceives of facts as abstract entities, the dichotomy is not particularly problematic.

We might run into problems when it comes to identifying which facts are reasons and which are not, but the very dichotomy itself seems sound. So on this, what I will refer to as, *thin approach* it is at least a possibility that some reasons qua facts will be (in some sense of the dichotomy) agent-neutral and some will be agent-relative.

The second approach, which I will comment on, the *thick account*, is, I believe, much less popular. However, as I shall argue, it would be better if we could understand the dichotomy even if we had this notion in mind. It would be preferable because it is better if our notion of a reason is consistent with a wider rather than narrower set of plausible

metaphysical views. But, more importantly, it would also be preferable because I think the thin approach, as I will explain in section 3 (the thin account), trivializes an interesting issue among practical philosophers. I will also argue that, as an additional drawback, the thin account is in one respect less appealing than its thick cousin. The thick account is not flawless, though. At the end, I will discuss a couple of objections to it.

Before we continue, a caveat is in place. As already mentioned, the dichotomy has been understood in a number of ways. I will set out from what is sometimes referred to as the *essentialist sense* of the distinction.

This is what I take to be Nagel's (but not necessarily e.g. Derek Parfit's) sense¹. Nagel draws the distinction in terms of what does and what does not *essentially* refer to a particular person, the owner of the reason. Various other writers have employed similar accounts.

Three notions of fact

In order to understand the reason-qua-fact idea, we need to take a short detour and formulate some idea about what kind of statements express these reasons. It is not obvious, of course, just how we should formulate a reason *statement*, i.e. a statement expressing a fact that constitutes a reason. Moreover, since there are different kinds of fact, and we should be open to the possibility that they can all be reasons, most likely there will be different kinds of reason-statement. To simplify matters, I suggest that we simply understand them as expressing what may figure in a *that-clause*. I therefore propose we set out from the following examples:

AN: A person *a* is drowning in circumstances *S*.

AR: My daughter is drowning in circumstances *S*.

Both statements refer to what we typically would refer to in order to answer questions of the form "Why should I/you Φ (e.g., jump into the water)?"

¹ Originally Nagel (1970) referred to the distinction as one between subjective and objective reasons. The terminology 'agent-neutral and agent-relative' was introduced by Derek Parfit in (1979). See Thomas Nagel (1986); especially (152-3).

Presumably, both AR and AN, describe only one aspect of a more concrete state of affairs². Let us for the moment focus on the one expressed by AR. Other features of this state of affairs would be, for example, *that the person (referred to in AR as drowning) has curly hair, lives in Barcelona, and desires to be rescued in S*. Needless to say, the list of features can easily be made very long. In fact, there seems to be no end to how many different descriptions we can give of a concrete entity.

Understanding the statements in this way raises the question just what *is the reason-constitutive element in AR (and AN)?* What precisely is it that constitutes the reason? Prima facie, only a few options seem reasonable: it might be the thin feature(s)³ explicitly expressed by the statement —*that my daughter is drowning in S*— or it might be the thicker state of affairs of which *that my daughter is drowning in S* is only one aspect out of many?

What I will refer to as the *thick account of reasons* suggests that reason statements refer to concrete entities; i.e., it is thick entities that are the reason-constitutive parts rather than the thin features. The state of affairs is, then, concrete in the sense that it contains a number of features that may be abstracted from these entities⁴.

Some of these features are more salient than others, and what e.g. AN and AR express is these salient features rather than the whole picture. On the thick approach the propositional element in a reason statement is like the tip of an iceberg; it displays only a tiny fraction of the whole entity.

The main challenger is the thin *account*. This approach, which I believe is more widely accepted, comes in two versions. However, to begin with it is important to notice that it should not be understood as denying, e.g. that the statement *that my daughter is drowning in S* describes a feature of a thicker state of affairs. But, for our purposes, the important thing to stress is that according to the thin *account*, the

² Does it have to be a unique state of affairs? Not necessarily. I am at least inclined to share the view of those who believe that the “closest world” in fact refers to a set of possible worlds that all share the property that they are equally close in resemblance to the actual world. However, for my purpose in what follows nothing important hinges on whether the thin feature is considered as an aspect of one unique closest possible world or of many equally close possible worlds.

³ The thin account may, of course, take a reason to be constituted by two or more features.

⁴ I would also add that in contrast to abstract features, these concrete entities may consist of objects in a wide sense (persons, things, etc.).

reason-constitutive element is *only* effectuated by the thin feature, and not by the thick state of affairs. So whether or not there is some other more concrete state of affairs involved, what makes AR or AN into a true reason statement is an *aspect* or a conjunction of aspects of a state of affairs rather than the concrete state of affairs.

One thing needs to be made clear. In this work I distinguish between “state of affairs” and “the obtaining of a state of affairs”. This allows me to regard states of affairs as one kind of entity and obtaining of states of affairs as a second kind of entity. Since there are, say, no mermaids, there is no obtaining of the state of affairs that mermaids live in the sea.

Moreover, since we are supposing that AN and AR express something true, they must therefore be understood as referring to an *obtaining* of the state of affairs rather than to a state of affairs. In what follows I will sometimes speak about these obtaining as concrete state of affairs.

According to the second, and I believe, more widespread version of the thin account, we should rather understand reasons as being the *content of possible judgements that are true*.

This suggestion sets out from a different perspective than the one expressed in the first version. At least, it would be possible (many would even say reasonable) to regard the first version of the thin account as being about the truth-maker of the proposition which the second version considers to be the reason-constitutive entity.

In what follows, to simplify matters, with the “thin account” I will have the latter, propositional version in mind. These accounts (as well as a combination of both) merit different treatments, and so my treatment of the thin account will necessarily be incomplete.

However, discussing all accounts would have made this paper considerably longer. Moreover, since, I believe, the propositional approach is the more popular, I have chosen to discuss it rather than the “feature approach”. (Later on I will briefly point out why the latter approach is problematic).

The thin account

There are advantages as well as disadvantages with the propositional account of the thin account. The most obvious advantage, I suppose, is that it certainly allows us to draw the distinction between agent-neutral and agent-relative reason *statements* (in the essentialist sense). Since the reason dichotomy is intuitively plausible, we should therefore welcome it. Moreover, it allows us to draw the distinction in a fairly simple way: We all agree that any statement contains either at least one indexical (e.g., 'I', 'my', 'he') or no indexical at all.

Hence, it would be absurd to deny that at least some possible proposition essentially refers to the person who issued one of these statements and thus endorsed the proposition. And since it is safe to say that at least some such propositions are true, we have cleared the ground for the existence of agent-relative reasons for a view that takes reasons to be thin entities. Some of these true, essentially referring, propositions will (it would at least seem so), be reasons.

In a structurally similar way, we can reason about statements and propositions that are true but are not essentially referring to the person whose reason they express. Some of these are suitably agent-neutral reason propositions. So if the true statement (proposition) refers essentially to the agent for whom the proposition is a reason, then it is an agent-relative reason; otherwise it is an agent-neutral reason.

This, then, is obviously a plus. But, in my view, it is a somewhat stained advantage. The discussion whether there are any agent-neutral reasons or, for that matter, any agent-relative ones, does not only turn on whether we are able to formulate true statements (and hence express propositions) that contain or do not contain ineliminable references. Certainly, the controversy concerns an evaluative or normative matter; it concerns, among other things, whether an ineliminable reference to an agent gives rise to a particular kind of reason, viz. an agent-relative reason. It does not concern whether a certain true statement or proposition involves such an ineliminable reference. That is not an evaluative issue.

Therefore, to suggest that we can draw a distinction between agent-neutral and agent-relative reasons if we acknowledge that some statements refer essentially to the agent and some do not, might clear

the ground for the distinction but it seems nonetheless to miss out on something important. It trivializes an interesting debate, which is unfortunate. Later on, in section 5 (problems for the thin account), I will substantiate this claim in greater detail. However, for the time being, a couple of other possible advantages ought to be mentioned.

One thing that might be talking in favor of the thin account is the following: It is generally acknowledged that reasons play or at least should be able to play an important part in our lives. However, if reasons are concrete obtaining of state of affairs, clearly they cannot play the sort of direct “in the head” role as a thin propositional entity appears to play. If reasons are the content of possible judgements that are true it becomes easier to understand how reasons could move us to actions, and hence play a more direct role in our lives. Or so the argument goes.

However, this “In-the-head advantage” is actually not much of an advantage⁵. Since we are talking about reasons as the content of possible thoughts that are *true*, then, if we have at least a classical view of truth in mind, reasons obviously will be entities with extrinsic features. They are therefore necessarily entities that relate to that which makes the thought true. Being, then, essentially a relational entity to something outside the agent (or at least to something outside the thought that *p*), suggest that the “in-the-head-advantage” is not really an advantage after all. What is in the head is more accurately (or, at any rate, just as correctly) described as *the grasping of the fact or feature, but not the fact or the feature itself*.

An advocate of the thick account can therefore rightly point out that she too can agree that the grasping, e.g. the believing that Anna is drowning, goes on in the head. She will just deny that this constitutes the reason, and that this should be recognized as well by the supporters of the thin account. After all, in their view, reasons encompass relate, one of which is external to the thought that *p*.

Notice, by the way, that the above objection is, I believe, consistent with the substantial idea that the only facts that constitute reasons are facts about the agent’s beliefs and desires. Unless we want to say that the facts are in the head of persons, the above point stands unaffected by

⁵ Observe that if we understand the thin account to be about features rather than the proposition, the advantage disappears.

the idea that only facts about the agent's beliefs and desires constitute reasons for her. Or more guardedly put: the point is not that facts are not in the head, but rather that it is at least as plausible to say that they are not in the head as it is to say that they are in the head⁶.

One thing that also perhaps talks in favor of the thin account is the fact that some reasons are hard to understand as anything but thin entities. Consider the following way of talking about reasons, which I borrow from Schroeder (11): suppose someone asked us why there are no circles with some points further away from their centre than others. We might reply the reason is that being a circle means having all of the points equidistant from the centre.

The "because fact" that we allude to here is more easily understood, I suppose, as a proposition (period); it is at least not clear to me what the concrete obtaining of a state of affairs would be in such a case. Its truth-maker seems to be some thin rather than concrete entity. We need not get into the complex matter of what makes such propositions true. But whatever it is, it does seem like a thin rather than a concrete object. Suppose this is correct. In that case, should we perhaps count this as evidence for the hypothesis that reasons are thin entities? Yes, I suppose; it is some evidence.

However, we should not be impressed by this sort of quick reasoning, which, unless qualified, simply boils down to a hasty generalization. Perhaps some reasons are like the above examples, but I very much doubt that all practical reasons are.

Core reason elements

In section 5 (problems for the thin account), I will bring up what I take to be two major disadvantages with the thin account. First, however, we need to get a better grip on what reasons actually involve according to the thin account.

⁶Moreover, a view that made the feasibility of our reason notion or reason dichotomy dependent on taking sides on this matter would therefore not be particularly appealing. This, I believe, cautious point will nevertheless not sit well with everyone. But whether or not we are comfortable about discussing where to locate facts, it would be unfortunate if the dichotomy were only applicable by advocates of the idea that only facts about the agent's desires and beliefs are reasons.

It is customary to look upon propositions involving an indexical (typically e.g. 'I' or 'my') as if they are someone's proposition, i.e., they are what a certain person expresses at a given time and at a given place. For the sake of uniformity, I suggest that we introduce this "owner" element whenever we talk about propositions. There is really no consensus on how propositions should be understood. However, in this context, I think this manoeuvre is rather inoffensive and non-committal. If we therefore keep this feature in mind, we may now *list* the key elements involved in a reason, according to the thin account, in the following way:

$$(i) \quad x [p] y (\Phi)$$

p is the true proposition; the function of the two kinds of parenthesis is to stress that their content is of a different kind than what is otherwise listed; x refers to the possible person entertaining the proposition, y refers to the person for whom the proposition is a reason, and Φ is the act which y has a reason to perform in virtue of the fact p . In some cases, however, x and y may as a matter of fact refer to one and the same person. For instance, this would be an example of an agent-relative reason, where TRR refers to me.

$$(ii) \quad \text{TRR [my dog is drowning] TRR (saving of the dog by TRR)}$$

Supposing (ii) is correct, we can now state why it expresses an agent-relative reason (on the essentialist account); the proposition that constitutes a reason for TRR contains an ineliminable reference to the person for whom it is a reason. There is a cross-reference between the person referred to as "my" and the person (the second TRR) for whom the proposition is a reason⁷.

On the other hand, the following would, then, be an example of an agent-neutral reason:

$$(iii) \quad x [\text{A dog is drowning}] \text{TRR (saving of the dog by TRR)}^8$$

⁷ (i) is a simplified picture that easily expands into the following: $xt [pt'] y \Phi t''$ here t refers to the time when x expresses the proposition, and t' refers to the time of the "propositional object", and t'' to when the reason applies to y , i.e. when y 's Φ -ing is normatively required. However, in this work I have had to simplify matters. E.g., the earlier suggestion that x in (i) might refer to the 'grammatical subject' has to be further specified, and it is also possible that p must somehow involve the act Φ . But since I believe a discussion of these matters would not substantially affect my arguments against the thin account, I have decided to leave out these qualifications.

⁸ But suppose x has TRR's dog in mind in (iii), and it is a reason for TRR. In that case, should we say that this still is an agent-neutral reason for TRR? I suppose so.

There is no cross-reference involved between TRR in (iii) and the proposition, and so, if this proposition (that a dog is drowning) constitutes a reason it constitutes an agent-neutral reason.

The above examples display the key elements involved, when something is a reason for a particular person, i.e. if we embrace the thin account.

So far, I have assumed that it is the true proposition p in (i) (i.e. $x [p] y (\Phi)$) that is the reason-constitutive element. However, this might be questioned. One might argue that the reason is not actually p but rather the whole substantial proposition p^* : *that* $x [p] y (\Phi)$, i.e. the entire content of (i).

However, I think there are good reasons to be wary about such a suggestion. First, recall that all I said about (i) was that it lists elements involved when some-thing is a reason. I did not say that (i) itself expresses a proposition. But perhaps there is indeed a proposition lurking in the vicinity. Maybe (i), in fact, expresses, say, that “ x endorses the true proposition p for y ” (or something to that effect), and that it is this fact, i.e. this true proposition, that constitutes a reason for y . This sounds somewhat awkward, however. It is not clear what it means to endorse a proposition *for* someone.

Maybe (i) should instead be read as expressing an even more complex proposition, viz. what is expressed by the following (N): “ x endorses the true proposition p , and p is R-related to y ” where R, then, specifies some kind of relation between (the endorsement of) p and the person for whom p is a reason. This (naturalistic) suggestion is not without its merits. Let us refer to it as the *Agent-endorsement account of reasons*, or AEAR. For instance, *explanatory* reasons are often said to refer to this sort of AEAR propositions, i.e., propositions like the one expressed in (N). When we want to understand why an agent acted as she did, we tend at least to invoke at some point a belief of the agent, especially a belief whose intentional content will have some motivational force for the agent⁹.

But although this might be one way to understand explanatory reasons, I think (N), or something to this effect, should be avoided as an analyses

⁹ A caveat is in place: In general AEAR would not suit as a motivational reason; this kind of reason should not be confused with explanatory reasons. What *motivates* an agent is seldom his belief that p ; it is rather that p .

of normative reasons. Admittedly, my reasons are mainly substantial. AEAR strikes me as simply unconvincing; why should we believe that of all the kinds of fact that might constitute a reason, it is only facts about what persons endorse (believe, think, etc.) that constitute reasons. That my dog is drowning would not on this account be a reason, only *that someone endorses* that my dog is drowning. There are further problems with this sort of view, but I think we may leave it and move on.

Problems for the thin account

The real worry with the thin account is that it locates the difference between agent-neutral and agent-relative reasons only in our descriptions of the world. I have already hinted at why I suspect this. The time has now come to spell out my worry in more detail. If substantiated it would certainly be, in my view, a serious drawback for the thin account.

Consider the (claim expressing the) proposition “Buster is drowning”. Suppose we accept the proposition as an agent-neutral reason maker for me, i.e. TRR. According to our schema, the following list, then, captures the core elements involved in an agent-neutral reason for TRR, assuming that Buster is a dog and that we have a reason to save a dog that is drowning:

- (iv) X [Buster is drowning!] TRR (Φ)

But, if (iv) is true, so is, *ex hypothesi*, the agent-relative (ii) TRR [my dog is drowning] TRR. That is, TRR’s proposition “my dog is drowning” constitutes an agent-relative reason for TRR. Both “Buster is drowning” and “my dog is drowning” refer, that is, to features of one and the same obtaining state of affairs (*viz.* one in which a particular dog that is called Buster and is owned by TRR is drowning). What is more, if (ii) and (iv) are true, so is, *ex hypothesi*, the following (where ARR refers to TRR’s daughter):

- (v) x (the dog of ARR’s father is drowning) TRR (Φ)

The thin account says, on the essentialist reading, that whether a true proposition is agent-neutral or not for TRR depends on whether the (corresponding statement expressing the) proposition p contains an

adequate proper name ((i.e. one that corresponds to the x in (i)), or an adequate indexical.¹⁰

However, to say something more substantial the thin account has to say, as a minimum, that there is a normative difference between (ii), (iii), (iv) and (v). It should normatively matter, in the sense that it should at least be feasible that only one of these might be a reason for *TRR*, or at least might be a stronger/weaker reason for *TRR* than the alternatives. Otherwise, the distinction would not be interesting or important. However, I fail to see that such a normative difference is detectable, and so I find the distinction trivialized.

Now, this point is far from obvious. Let me therefore explain how it resists the more obvious rejoinders. First, one might object that, depending on who will act on these propositions, it will certainly in some cases make a difference whether the person believed the proposition in for instance (iv) rather than the one in (v); so although we assume that both are made true by (or, if you like, are features of) the same concrete state of affairs, it is easily conceivable that grasping the proposition in (iv) might have a different effect on y than if she believed the proposition in (v).

For instance, suppose *TRR* *falsely* believes that *ARR* is not his child but that he correctly believes that Buster is his dog. In that case, it is quite expectable that, depending on whether the agent believes (iv) or (v), it will have a different impact on what she will do (or even take herself as having a reason to do or be, as a matter of fact, motivated to do). As is well known, the intentional context of a proposition is crucial when it comes to the effect that a proposition has on the agent's behavior.

I definitely do not want to deny this. And so, if we could reduce normative reasons to what goes on in the agent's mind, in terms of propositional attitudes, it is clear that we would have an interesting reason notion, and, in its train, a plausible reason dichotomy. But since,

¹⁰ That indexicals might cause trouble for the agent-relative/neutral dichotomy was made clear by Derek Parfit: "Even if you and I are trying to achieve some common aim, we may be in different causal situations. I may have reason to act in a way that promotes our common aim, but you may have no such reason since you may be unable to act in this way. Since even agent-neutral reasons are, in this sense, agent-relative, this sense is irrelevant to our discussion" (143). However, just what Parfit commits to here is a matter I leave open. His understanding of the dichotomy does not quite coincide with the essentialist reading. For more on this, see Rønnow-Rasmussen (2011).

I think it is safe to say, the normative still resists this kind of reduction, this sort of rejoinder currently misses its target.

It might be objected that this objection to the thin account rests on an unwarranted assumption. At least, it seems as if I assume here that what makes the proposition true is some concrete state of affairs. But, it might be argued, it is in effect not some concrete state but rather only the abstract feature(s) that have this role. It could accordingly be argued that since these features, corresponding to, say, what makes (i)-(v) true, are different, we should also say that (i)-(v) express different reasons. And since these features may in different possible situations call for different sorts of actions, we should conclude that this sort of view handles the objection from above.

I am not convinced by the above reasoning, though. Suppose we agree that it is different features, *period* (rather than *features of a concrete entity*) that make, say, (iii), (iv) and (v) true, and that we accordingly should say that these three true propositions are or at least refer to different reasons. This in itself does not meet the challenge; it still needs to be argued that individuating reasons in this way is normatively interesting.

This shouldn't simply be assumed. In fact, when these features make the propositions true in the same world, at the same time, it still appears puzzling to me how one could say that it normatively matters in the sense outlined above. I still do not see that it does.

These thin features (expressed by iii, iv and v), I take it, have relational properties to other features and in particular to each other in the case they are true in the same world, such that they will make true some further propositions that make it clear, with regard to the examples we are discussing, that, for instance, ARR is TRR's child. So I have a hard time seeing that invoking these features as separate reasons makes a normative difference.

One might object in a different way. I have said that at present it is still unclear whether normativity can be naturalistically reduced. Besides this qualification, I have left it open just what makes a fact into something normative. But suppose we endorsed an account along the lines of John Broome's (2004) interesting suggestion that a normative reason for Φ -ing is what *explains why the agent ought to Φ* . In that case, it seems as if it would make a difference whether we offered the explanation in terms

of, say, “the dog of ARR’s father is...” rather than in terms of “TRR’s dog is...” This reply works, I think. If normativity is understood in this way, then it is conceivable that it normatively matters whether we (iii), (iv) or (iv) is a reason.

However, although appealing, Broome’s suggestion is not an obvious one. What is an explanation in one context need not be so in another; explanation is a success notion in a way that the notion of a reason is not. Intuitively, it is at least conceivable that I could have a reason for Φ -ing, even if every attempt to explain why I ought to Φ failed. Perhaps we can get round this sort of intuitive objection. At present I do not see how, though.

The second problem with the thin account is that it is aesthetically less appealing than the thick account. Recall the examples from above (ii, iii, iv, v).

The unattractiveness derives from the following: since it is easy to imagine that one and the same concrete entity displays endless variants of the aforementioned features, the thin account will seemingly have to say about our example that the agent faces an endless list of reasons for saving the dog.

This strikes me as an odd thing to say. But perhaps accepting this might, after all, be quite inoffensive. Perhaps we might just say that there is only one reason but lots of versions of it. How inoffensive such claims are depends, I suppose, on whether we take each of these reasons (or versions of reasons) to have normative strength, and, in particular, whether we think these strengths somehow add up.

The latter seems hard to justify, however. After all, it would have some rather counter-intuitive implications. If each of these reasons carries a certain normative strength, we appear to face the strange conclusion that at any given time that an agent has a reason to perform an act, he or she has infinitely strong reasons to do so – strength here being interpreted as the sum of the strength of each reason. This also strikes me as odd.

An advocate of the thin account is therefore well-advised to deny that all these reasons somehow add up in strength. But, of course, in that case, it seems reasonable to ask why we should in the first place crowd the canvass with countless reasons (it gives us a lot of unnecessary trouble,

such as finding out which true propositions carry normative strength that may add up with the normative strength of some other proposition); it seems better (i.e., gives a cleaner picture) to say that there is actually only one reason (or, in the last example, two pro-tanto reasons), viz. one that is constituted by a concrete fact that can be described in numerous correct ways.

The thick account

The idea that reasons are concrete is, in one respect, a simpler model; it involves fewer elements, as the following list shows:

(vi) $[O]y (\Phi)$

Here O refers to the obtaining of a state of affairs. There is no mention of any proposition or -what is more important- the person whose proposition it is. However, the thick account introduces the complicated notion of an *obtaining*, i.e. a concrete entity. Needless to say, this idea is much less transparent than what could be wished for.

But it is important not to overlook that whatever fault we take this to be for the thick account; it will most likely also be a flaw for the thin account. In any case, it will not be a simple and straightforward matter to do without the distinction between a state of affairs and its realization (i.e. the obtaining of a state of affairs).

So it is far from clear that the thin account can profit from whatever flaws we detect in the thick account's core notion. But the thick account can, on the other hand, take advantage of the faults we detect in the thin account due to its idea that reasons are true propositions. Propositions play no constitutive part in the thick account.

As expected, the thick account faces in its turn some fairly common objections. Perhaps the most common one is the following: Although we often speak of facts, the idea that we can separate one fact from another is puzzling; there is actually only one concrete entity that can be described as the obtaining of a state of affairs, and it is the one which encompasses how the world is at a present point in time (or something close to this).

But this, then, seems to lead to the absurd idea that my reason for, say, saving a dog from dying, is constituted by this monstrous entity that encompasses all the features of the world at the given time that I have a reason to rescue the dog. The thick account should therefore be rejected.

There are different variants of this sort of *reductio ad absurdum* (and I must confess, not all the details are equally clear to me); basically they all involve two ideas that must not be conflated. The first one puts it to us that what constitutes a reason cannot be too complex or exceed a certain limit; e.g., it cannot incorporate too much or too many items. But, the argument goes; the problem with the thick account is that this is precisely what it does. The sheer size of what constitutes a reason, on the thick account, is so overwhelming that it should be evident that my reason for, say, saving the dog simply cannot be constituted by such a colossal entity.

The above “monstrosity objection” involves a related but nonetheless different idea. Even if we somehow accepted that there is no limit to the size of a reason-constitutive entity, the thick account is open to a perhaps more serious objection. The sort of entity that it suggests is constitutive of a reason would simply involve features that it would be implausible to include in the supervenience base of the reason. Whatever it is that makes it into a reason that the dog should be saved, it certainly does not involve facts about, say, people and objects on the other side of the planet. This, as we might call it, “too inclusive” objection seems to me to be more serious than the former one. However, to begin with, let me briefly comment on the former objection.

I think there are mainly two explanations close at hand for why anyone would suggest the monstrosity objection. The first stresses something that we have already commented on, namely that reasons should be in some sense accessible to us. Simplifying, the point can be expressed as follows: *For something R to be a reason for a person x, R must at least be graspable by x*; otherwise *x* could not have acted for the reason R. Such a claim would most certainly have to be qualified. However, there is no need here to go into details. Essentially, what it says is that there is a necessary (conceptual) link between “R is a reason for x” and “x could act for a reason R”, i.e. “x could be motivated by R”.

The problem with this sort of view is that it simply does not seem to be true that there is such a necessity. Perhaps I am wrong about this, though. Once we qualify the circumstances, and disregard abnormal cases, irrational persons, etc., it will perhaps become manifest that there is such a tight relation between ‘being a reason’ and ‘being such that a person in circumstances C could act from it’. But again, this remains to be shown.

My suspicion is that reasons need not be graspable. However, since I think (as will become clear in a moment) the thick account actually is consistent with “the graspable requirement”, for reasons I will get to in a moment, I will not argue for this here¹¹.

The second explanation is quite different. It sets out from the distinction between a *pro-tanto*, or contributory reason, which it takes to be constituted by some aspect(s) of the world, and a *complete* reason, which it understands as what would be called for if we took “everything” into consideration. It then argues that the thick account somehow assumes that we can have a “complete reason”. But, as the argument has it, the notion of a complete reason is not coherent, and so we should give up the idea that reasons must be understood as thick entities.

For sure, if we cannot find a reasonable way of understanding “complete reason”, we are handed a strong incentive for giving up the thick account. However, I am not sure that the notion of a complete reason is really inconsistent. For reasons I will get to in a moment, I think it would be unwise and certainly premature to set this idea aside as being incoherent.

The *too inclusive* objection carries greater force than the two above objections. The gist of it is that we should reject the thick account because otherwise we would end up with a set of subjacent features that simply cannot be what makes something into a reason. For instance, *ex hypothesi*, in this world where Buster is drowning, there are numerous features that are irrelevant. We do not want to have to say

¹¹ There is a weaker condition, which seems more plausible, viz., that the parts of a reason should be accessible to the agent. Although the whole reason might be a real mouthful, impossible to grasp, each of its constitutive parts might be graspable. This suggestion, which I owe to Wlodek Rabinowicz (personal communication), does seem plausible. Moreover, it is quite accessible to advocates of the thick account.

that my reason for saving the dog involves, among other things, facts about, say, the number of leaves on one particular tree 10,000 miles away from the dog and me. This would definitely be absurd to accept.

Notwithstanding this, I think the aforementioned objection can be discarded. One way to do so would be to admit that the supervenience base includes a somewhat peculiar kind of fact, viz. a fact about facts. If what calls for me acting in a certain circumstance are not only certain positive features but also the absence of countervailing features, we have a response to the “too inclusive objection”.

Thus, what makes something into a complete reason are certain salient features, together with the peculiar feature that any remaining feature is irrelevant. We do not have to give up the idea that among the reason-making characteristics of the world, we should only include those that are relevant for the explanation of why the agent ought to Φ . It is just that one of these relevant features is the feature that there are no other relevant features. So if at a given time the facts are such that certain features positively call for some action Φ by me, and it is the case that there are no other features that would discard this call, then I have a complete normative reason for acting.

I am inclined to accept this sort of “complete reason”. The drawback is that it seems that in order to explain normative reasons, we will have to do so in terms of what appear like normative notions, namely “relevance/irrelevance”. Perhaps this should make us look elsewhere¹².

However, I am not yet clear whether this circularity is really devastating for the analysis. Perhaps it is wide enough to be illuminating.

I have outlined two ways of understanding the idea that reasons are (constituted by) facts. I have also argued that it would be preferable if we could understand the agent-neutral and agent-relative reasons dichotomy in such a way that it is consistent with both these views. The idea behind this contention is that the thin account is inferior to

¹² We might avoid the circularity by suggesting that ‘relevant/irrelevant features’ refers to an evaluative but not necessarily normative feature. For instance, it might be claimed that the distinction is not really referring to what are reason-making features. Rather, a relevant/irrelevant feature might be described as one that is e.g. good/bad at motivating us, or at least at explaining why you should do a certain act, or, what calls for an action in the most effective way. Just how viable this suggestion is needs to be further examined. However, it is clearly not an option for advocates of so-called Buck-passing accounts or Fitting attitudes analyses of value.

the thicker account in at least some aspects. Of course, this conclusion does not suggest that the reason dichotomy is in the clear. The thick account might lead to complications for the distinction in a way that the thin account avoids. In fact, I believe this to be the case. However, this is not something that I will argue here.

BIBLIOGRAPHICAL REFERENCES

Broome, John. "Reasons". *Reasons and value: Themes from the moral philosophy of Joseph Raz*. Ed. R. J. Wallace. Oxford: Oxford University Press, 2004. 28-55. Print.

Nagel, Thomas. *The view from nowhere*. New York: Oxford University Press, 1986. Print.

Parfit, Derek. "Prudence, morality and the prisoner's dilemma". *Proceedings of the British Academy*, 1979: 539-564. Print.

Rønnow-Rasmussen, Toni. "Normative reasons and the agent-neutral/relative dichotomy". *Philosophia*, 2009: 227-243. Print.

---. *Personal value*. Oxford: Oxford University Press, 2011. Print.

Schroeder, Mark. *Slaves of the passion*. Oxford: Oxford University Press, 2007. Print.

Svensson, Frans and Rysiek Sliwinski, Ed. *Neither/Nor-philosophical papers dedicated to Erik Carlson on the Occasion of His Fiftieth Birthday*. Uppsala: Philosophical Studies, 2011. 243-257. Print.

RECONCEIVING PHILOSOPHY OF RELIGION

RECONCIBIENDO LA FILOSOFÍA DE LA RELIGIÓN

PAUL K. MOSER

Loyola University Chicago, USA. p.moser@luc.edu

RECIBIDO EL 2 DE MARZO DE 2012 Y APROBADO EL 18 DE MAYO DE 2012

RESUMEN ABSTRACT

Los filósofos han trabajado durante mucho tiempo con concepciones de Dios, inadecuadas para representar a un Dios genuinamente digno de adoración. Una deficiencia importante es la omisión de la noción de severidad divina, apropiada para la idea de un Dios digno de adoración. Como resultado, muchos filósofos tienen expectativas equivocadas sobre Dios, es decir, expectativas que no concuerdan con lo que serían los propósitos relevantes para Él, si Dios existiera. Estos últimos propósitos incluyen aquello que Dios busca lograr cuando revela a los seres humanos (la evidencia de) su realidad y voluntad. Las expectativas equivocadas de Dios nos pueden llevar a buscar pruebas de su existencia en todos los lugares equivocados.

El antídoto necesario requiere una reconsideración cuidadosa de nuestras expectativas sobre Dios, y nos capacita para acercarnos a una epistemología de la religión, de una manera que le hace justicia a la idea de un Dios digno de adoración. El artículo sostiene que la evidencia disponible a los seres humanos de un Dios digno de adoración, no sería para meros espectadores, sino que buscaría desafiar la voluntad de los seres humanos para cooperar con la voluntad perfecta de Dios, como sucede en el caso del desafío divino de Getsemaní.

Philosophers have long worked with conceptions of God inadequate to a God genuinely worthy of worship. A key inadequacy is the omission of a notion of divine severity appropriate to the idea of a God worthy of worship. As a result, many philosophers have misguided expectations for God, that is, expectations that fail to match what would be God's relevant purposes, if God exists. The latter purposes include what God aims to achieve in revealing to humans (the evidence of) God's reality and will. Misguided expectations for God can leave one looking for evidence for God in all the wrong places.

The needed antidote calls for a careful reconsideration of our expectations for God, and enables us to approach religious epistemology in a way that does justice to the idea of a God worthy of worship. The article contends that the evidence available to humans from a God worthy of worship would not be for mere spectators, but instead would seek to challenge the will of humans to cooperate with God's perfect will, as in the case of the divine challenge from Gethsemane.

PALABRAS CLAVE KEY WORDS

Concepto de Dios, evidencia, expectativas de Dios, fe, Getsemaní, Dios, severidad divina, voluntad.

Concept of God, evidence, expectations of God, faith, Gethsemane, God, severity, volition.

Expectations of God

Our expectations of God, if God exists, can get in the way of our receiving salient evidence of God. Perhaps we assume that God would have certain obligations to us, even by way of giving us clear evidence, and when those obligations are not met, we discredit God, including God's existence. This is a fast track to atheism or at least agnosticism. We need, however, to take stock of which expectations of God are fitting and which are not, given what would be God's perfect moral character and will. If God lacked a perfect moral character and will, then God would not be worthy of worship, and hence would not satisfy the supreme title "God". We then might have a god, but not God.

Perhaps God is not casual but actually severe regarding available evidence of God, owing to God's intense concern for the realization of divine righteous love (*agapē*), including its free, unearned reception and dissemination among humans. Perhaps the latter concern stems from God's aim to extend, without coercion, lasting life with God to humans, even humans who have failed by the standard of divine *agapē*.

God's severe commitment to that goal could figure in God's making human life difficult, or severe, for the sake of encouraging humans, without coercion, to enter into a cooperative good life with God. This severe God will not sacrifice a human soul to preserve human bodily comfort.

The *Oxford English Dictionary* (1989) offers this main definition of "severity": "strictness or sternness in dealing with others; stern or rigorous disposition or behaviour; rigour in treatment, discipline, punishment, or the like". This definition does *not* entail moral badness or evil, or any moral deficiency for that matter, contrary to some less prominent uses of "severity". The severity in question, however, does involve discomfort, anxiety, stress, rigor, or insecurity.

Being perfectly and severely active in gracious righteousness, God would oppose whatever obstructs perfect righteousness among agents. This opposition would be wisely intentional, and not impulsive or irrational. It could allow, however, for some short-term unrighteousness for the sake of long-term righteousness. So, God's opposition to certain occurrences of unrighteousness, such as unjust human warfare, could be eventual rather than immediate.

God could allow some episodes of unrighteousness to persist for a while in order to have them culminate and be seen as harmful by a wide audience. A sound morality would not require God to oppose all unrighteousness immediately if a purpose of greater righteousness would be served only by means of postponing opposition. Opposition to unrighteousness, then, does not entail immediate opposition to all unrighteousness¹.

Severity in God

The New Testament offers some ascriptions of severity to God, and these ascriptions can illuminate how God can be oppositional. In Luke's version of the Parable of the Pounds, Jesus attributes the following statement to a man who functions as God's approved representative and thus images Jesus himself and even God: "You knew that I was a severe (*austēros*) man, taking up what I lay down and reaping what I did not sow" (Luke 19, 22) (cf. Matthew 25, 24).

Likewise, Paul remarks as follows, in connection with the divine offer of mercy to humans coupled with the divine judgment on human resistance: "See then the kindness and the severity (*apotomian*) of God" (Romans 11, 22). Evidently, then, some of *The New Testament* writers would propose that divine love has a certain severity about it.

We can use the previous OED definition to clarify the talk of divine severity in Luke and Paul. We begin with the idea that divine severity involves "strictness or sternness in dealing with others; stern or rigorous disposition or behaviour; rigour in treatment, discipline, punishment, or the like". We need to clarify this talk of severity to capture a fruitful conception of God often ignored by philosophers, theologians, and other reflective people. This neglected conception prevents us from thinking of God as akin to a doting grandparent or a celestial Santa Claus figure. The neglected conception rightly preserves genuine moral *gravitas* in God and in God's dealings with humans.

Since the time of Dietrich Bonhoeffer's work on discipleship (1937), many people have been wary of "cheap grace" from God. Such grace would not challenge its recipients to undergo profound transformation toward

¹ The relevant notion of "greater" righteousness is complicated, of course, but we need not digress.

righteousness corresponding to God's moral character. According to Paul, in contrast, the divine gift of grace includes the gift of righteousness and is therefore profoundly morally transformative (Romans 5, 17).

In Paul's view, God's redemptive plan is that "grace might reign through [the gift of] righteousness to eternal life" (Romans 5, 21). Given this key linking of grace and righteousness, we should be suspicious of any "cheap theism" that either makes God morally lax toward unrighteousness or divorces divine grace from divinely empowered righteousness among humans. If severity is a divine meta-attribute that applies to the divine attribute of righteousness (among other divine attributes), such cheap theism will misrepresent God's moral character.

Philosophers of religion and theologians speak often of divine "love", but they talk much less frequently, if at all, of divine *severity* toward unrighteousness. The result is a conception of a God who is alleged to "loving", in some sense, but who is anything but a "consuming fire" of moral righteousness (cf. Hebrews 12, 29 Deuteronomy 4, 24).

This conception sacrifices divine righteousness for a kind of love that is soft on unrighteousness. We shall ask whether this conception fits with the moral character of a God worthy of worship. Perhaps divine love involving worthiness of worship is severely righteous in a way that opposes unrighteousness of any kind. If so, God would be actively *oppositional* and even *severely* oppositional toward unrighteousness. We can find a divine redemptive goal of the moral transformation of humans in such oppositional behavior, which bears importantly on human knowledge of God.

Many biblical passages yield a striking portrait of a severely oppositional God. In the Hebrew bible, God is portrayed as inherently jealous: "you shall worship no other god, because the Lord, whose name is Jealous, is a jealous God" (Exodus 34, 14) (cf. Deuteronomy 6, 14-5).

This divine jealousy differs from, and is even incompatible with, selfish human jealousy. The jealousy in Israel's God aims to protect the people of Israel in their relationship with God by opposing all idolatrous substitutes, visible or invisible. Accordingly, this God announces to the wayward people of Israel: "I will chastise you in just measure, and I will by no means leave you unpunished" (Jeremiah 30, 11) (cf. Exodus 20, 5).

God intends such just chastisement to be corrective and life-giving, according to Jeremiah and other Hebrew prophets; it is not an end in itself. Even so, as punishment, it can be retributive, or compensatory, in virtue of penalizing behavior opposed to God's character and life. The same holds for the notorious divine wrath that motivates the prophesied punishment of people disobeying God; it, too, can offer a penalty as well as a purported correction. A penalty could include, for instance, severe human discomfort or stress intended to draw attention to a needed correction among wayward humans.

The New Testament portrays God as oppositional toward unrighteousness in a number of ways. Paul writes: "The wrath of God is revealed from heaven against all ungodliness and wickedness of those who by their wickedness suppress the truth" (Romans 1, 18). He also remarks:

The creation was subjected to futility, not of its own will but by the will of the one [namely, God] who subjected it, in hope that the creation itself will be set free from its bondage to decay and will obtain the freedom of the glory of the children of God. (Romans 8, 20-1)

The divine aim, or hope, is to free people from bondage to what does not give a life of freedom with God. Pursuing this aim, God subjects creation to futility, according to Paul, in order to manifest and to overcome the inadequacy of the created world for the life of freedom needed by humans. *The New Testament* message implies that God alone will emerge as adequate for providing such life.

Paul suggests that God intends tribulation in human life to produce human character and hope agreeable to God's morally perfect character and life (Romans 5, 3-5). Human tribulation, then, is not portrayed as an end in itself.

Another way to oppose unrighteousness includes God's supplying receptive humans with a special kind of power. This is the power to "put to death the [disobedient] deeds of the body" that alienate people from the life of freedom with God (Romans 8, 13).

In this view, humans need help, particularly helpful power, from God to overcome the world's pull toward unrighteousness. The helpful power would be an antidote, at least for willing humans, to unrighteousness as

disobedience toward God. Divine opposition to human unrighteousness, accordingly, would aim for a human life of freedom reconciled to God.

Divine opposition to human unrighteousness has various manifestations and resists any simple characterization. It includes the divine severity of subjecting creation to futility, or frustration, in order to liberate people from their bondage to decay in inadequate sources of freedom and security. Such divine opposition occurs in at least some human suffering, dying, and death, but an important qualification is needed. God could subject creation to futility, for redemptive purposes, without *directly* bringing about every instance of futility in creation.

God could create free agents who freely bring about some of the world's futility and are therefore causally and morally responsible for it. This option can raise serious problems for a characterization of God as successful in realizing God's will at every point. The latter characterization would have to face the reality of the power of created free agents, if they exist, to frustrate God's will. Part of God's noncoercive "providence", then, may be to allow other agents to exercise control, even harmful control, over parts of creation.

God would not have to be the direct cause of all that occurs in creation, because God could allow some opposition to unrighteousness to arise from the causal powers of created agents. In that case, God's *permissive will* would be operating, even if God's *executive will* would not. God's allowing an action by another agent does not entail God's causing, performing, recommending, or approving that action.

A memorable case of divine severity emerges in the crucifixion of Jesus as a self-avowed emissary for God. Aside from the theological significance of his crucifixion, the human treatment of Jesus just before his death was remarkably severe. Some theologians assign this severe treatment ultimately to God, and not just to the Roman officials and soldiers, on the ground that God chose to punish Jesus to save humans from the just desert of their sins.

This interpretation is highly controversial, partly because it lacks clear support in *The New Testament*, and, in any case, it raises serious questions about divine justice. Nonetheless, God's *allowing* the crucifixion of Jesus is severe by any ordinary standard of severe permission, even if certain Roman officials and soldiers were causally and morally responsible for the harsh punishment involved.

The God of the Hebrew bible sometimes judges people by withdrawing the divine presence from them. God reports to Moses:

My anger will be kindled against them [the people of Israel] in that day. I will forsake them and hide my face from them; they will become easy prey, and many terrible troubles will come upon them. (Deuteronomy 31, 17)

In such a case, divine hiding involves severe divine punishment, even if the recipients are unaware of the punishment².

Some commentators apply the idea of divine hiding to Jesus's cry of abandonment on the cross: "My God, my God, why have you forsaken me?" (Mark 15, 34) (cf. Psalms 22, 1 Matthew 27, 46). The idea is that, in order to condemn human sin, God withdrew divine presence and fellowship from Jesus, as the atoning representative of humans, at the time of his troubled cry to God.

This is arguably a plausible way to unpack Paul's following remark:

God has done what the law, weakened by the flesh, could not do: by sending his own Son in the likeness of sinful flesh, and to deal with sin, he condemned sin in the flesh. (Romans 8, 3)

It does not follow, however, that in condemning "sin in the flesh" God condemned Jesus himself. The message of Paul is that God ratified Jesus and his obedient life by resurrecting him, but only after allowing a severe death by crucifixion. This message offers Jesus as a personal model of human dying into lasting life with God.

Three questions can illuminate the potential ways of God toward humans. These questions may be called *expectation-evoking*, because they are helpful in eliciting sound expectations regarding God and God's ways of intervention in human lives. Inquiry about God would gain enhanced precision if more attention were given to such questions. Philosophers and theologians have given inadequate attention to expectation-evoking questions about God in connection with the problem of divine severity. This paper begins to correct that deficiency.

² It does not follow that all divine hiding entails God's punishment of humans; see Moser (2008).

Personalism

First, how should we expect God to relate to humans if God aims to be known directly by them *as a personal agent*, and not as a principle, an idea, or an impersonal cause? If God is worthy of worship, then God is a morally perfect, perfectly loving agent and therefore is personal (rather than nonpersonal). In virtue of being personal, God would be an *intentional* agent, with purposes and plans, and God's being personal could be reflected, if imperfectly, in human persons.

Accordingly, God could not be replaced without loss by a principle, an idea, or an impersonal cause. Any such replacement would dispense with perfect *agapē*, for instance, because only a personal agent can offer perfect *agapē*. Principles, ideas, and impersonal causes do not love anything, or perform any intentional actions, for that matter.

Direct knowledge of God as a personal agent with a will would serve an important redemptive purpose. It would give humans a direct apprehension of God that could not be supplied by principles, ideas, or impersonal causes. In doing so, it would reveal God's irreducibly personal character in a way that principles, ideas, or impersonal causes could not. Such direct knowledge thereby would convey an important feature of God's actual character to humans, and could be an experiential avenue to awareness of God as worthy of worship and perfectly loving.

Aiming to be known directly as a personal agent, God would value *interpersonal interaction* of a direct, second-person sort. Such interpersonality is *de re*, or more accurately *de te* (from the Latin "tu" = "you"), involving the direct acquaintance of one personal agent with another personal agent in the second person, beyond any *de dicto* (conceptual or notional) relation involving ideas or principles. This kind of acquaintance between God and humans is not bodily acquaintance, because God has no physical body.

The most plausible human context for direct acquaintance with God is human *conscience*. This is the psychological place where a human could directly know, and be known, together with God (see the etymology of "*con + scientia*") as God calls a person (sometimes, to account) in the second person, as *you*. This proposal fits with Paul's suggestion that in human conscience God bears witness to the divine moral character as represented in the law of God, thereby holding people accountable (Romans 2, 14-5).

It also fits with Paul's suggestion that his conscience can confirm something by God's Spirit (Romans 9, 1); (cf. 2 Corinthians 1, 12 5, 11). In this perspective, one's conscience is the inner place, involving one's spiritual "heart", where one can directly experience hearing from, being called by, or being taught by God (cf. John 6, 45) (Matthew 16, 17) (1 Corinthians 1, 9) (1 Thessalonians 3, 11 4, 7-9) (Hebrew 3, 7-15).

The role of human conscience in knowledge of God is widely neglected among contemporary philosophers, and this neglect can obscure the experiential reality of the God of Abraham, Isaac, Jacob, and Jesus. It also can minimize the crucial role of prayer in evidence-conferring interaction with this God. The latter role finds its perfect model in the prayer of Jesus to God in Gethsemane: "Father, not what I will, but what you will" (Mark 14, 36).

This paper contends that such a Gethsemane attitude toward the priority of God's will is central not only to cooperative life with God but also to human appropriation of evidence for God. This approach is neglected among philosophers of religion, past and present, but this paper begins to correct this neglect.

Our first expectation-invoking question leads to a subsidiary question: is there actually a Godward presence or call in human conscience? Human conscience can be insensitive, corrupted, and outright diabolical. We should be very skeptical, then, of Wittgenstein's following remark: "Certainly it is correct to say: Conscience is the voice of God" (75).

Even so, conscience is not thereby emptied of Godward value, just as perception does not lose its cognitive value given the misleading functions or uses of perception. Humans can treat conscience honestly or dishonestly and earnestly or indifferently. Dishonest or indifferent treatment of conscience does not undermine the representational value of conscience treated honestly or earnestly. Personal accountability thus can figure importantly in the handling of human conscience, including in its representing or not representing God to oneself. In addition, we should not expect conscience to coerce humans in its representing God, as if humans had no responsible interactive role toward God.

Many philosophers of religion seek principles that supply intricate human explanations of God's ways. This paper contends, however, that we have to deal primarily with a personal God who may not honor such principles (cf. 1 Corinthians 1, 18-25).

As a result, we should expect our characterizations of God in some areas to be less “cut-and-dried” and even less adequate than we might have wished. Even so, the paper contends that God can, and does, make Godself known via a receptive human conscience in ways that are much more personally challenging (even severe) and morally robust than the arguments of traditional natural theology.

We shall see that a distinctive kind of *personifying* evidence of God can be found in the personal moral character of a human agent, beyond mere propositions, claims, or arguments. For instance, John’s Gospel and Paul’s letters suggest that Jesus is the perfect human personifying evidence of God, and he is, of course, not a mere proposition, claim, or argument. As a personal agent, with definite intentions and plans, he can serve as personifying evidence for God, who also has definite intentions and plans of the same sort. Other humans, too, can become personifying evidence of God in virtue of their cooperatively receiving certain features of God’s moral character. In this approach, involving evidential personalism, persons can play a role in foundational evidence of God that cannot be reduced to mere propositions, claims, or arguments.

Looking for principles instead as ultimate, some philosophers of religion miss the distinctive point and value of evidential personalism regarding God. For instance, much of analytic philosophy of religion fails to consider some of the key features of Jewish–Christian theistic personalism, such as the bearing of Jesus himself on human knowledge of God.

This paper contends that we would do well not to sacrifice such personalism for any kind of deism, mere theism, or principle-based approach to God.

Deep transformation

Our second expectation-evoking question is this: how should we expect God to act in relation to humans if God aims to redeem them not just as thinkers but as *morally responsible volitional agents* who need a self-commitment to cooperation with God, for the sake of companionship with God? We may formulate the latter aim in terms of God’s aiming to win humans, themselves, and not just arguments with humans.

This divine aim would offer a gift of divine–human reconciliation that originates outside any human resources. It would include a person-to-

person (divine-to-human) call to be renewed at one's motivational center via a life of self-commitment to cooperation with God. Such a divine aim would feature the divine generosity of companionship offered to humans and, by way of an accountable response, the human cooperation demanded by God. Notoriously, the latter demand can lead to real severity in cases of human irresponsibility toward God.

From God's side, the companionship would include God's self-giving intervention in Christ on behalf of humans, whereby God seeks to self-identify and live with humans. Without such divine grace in action, many humans properly would doubt that they meet the standard set by God's own moral character, and they therefore would lack proper confidence regarding their acceptance by God. As a result, without divine grace in action, any self-commitment to God would be seriously deficient.

God's redemptive aim would be to give receptive humans a renewed volitional center of agency in companionship with God, including a renewed will to live agreeably and therefore unselfishly with God. The divine aim of human companionship with God would stem from the human need of such benefits as encouragement, chastisement, and specific personal guidance from God. It also would stem from the human need of seeking, including asking for, these benefits from God, as part of interpersonal communion. As a result, no mere principle, command, law, morality, argument, or human effort would accomplish this redemptive aim.

The transformation of humans in divine redemption would oppose moral self sufficiency in humans. Accordingly, it would oppose any presumption of humans being good *on their own*. Instead, it would aim for moral transformation *in human companionship with God*. For the good of humans, this would be reverent, submissive transformation anchored in the prayer offered by Jesus to God in Gethsemane: "Father ... not what I will, but what you will". Such transformation would be person-to-person, in a context where a human submits to God in companionship with God. The result would be significant human change via willing human participation in God's perfect moral character.

God's pursuit of the volitional transformation of humans would not be served by just abstract theoretical claims about God, such as claims about divine impassibility or divine omniscience. Claims of that sort would

invite philosophical and theological discussion, perhaps even endlessly, but they would not challenge the volitional center of an agent. They would not challenge an agent to make a commitment of self-sacrificial love to a God of redemptive grace. Concrete self-sacrificial actions and corresponding commands from God, in contrast, could infuse the needed challenge with motivational significance for humans.

Accordingly, the Christian message is that God's culminating revelation comes in the self-giving person and life of Jesus Christ, who manifests God's grace and wisdom, and not just in ideas about God.

We should ask what the relevant divine actions and commands would look like, given God's perfect character that includes divine severity. The separation of (evidence for) God's existence from (evidence for) God's perfect personal agency, as in much natural theology, invites serious problems and should therefore be avoided (see Moser 2010).

This paper identifies a role for human self-sacrifice, as submission of one's will to God, as part of a proper response to divine self-sacrifice. It argues that a philosophy of religion adequate to a God worthy of worship must award a key role to such human sacrifice to God.

Human boasting undone

Our third expectation-evoking question is this: how should we expect a God worthy of worship to intervene in human lives if this God seeks to undermine all selfish or otherwise misguided human boasting? One plausible answer is straightforward: by an undeserved gift of divine grace that displaces any such boasting. This displacement of human boasting would challenge misguided human self-pride on various fronts, including in connection with morality, knowledge, and wisdom. Such pride involves one's exaggerated assessment of oneself or of one's contributions or achievements, such as when one self-confidently takes credit for something (for instance, in morality, knowledge, or wisdom) where such credit is manifestly not due.

A familiar example includes a human's self-confidently taking credit for what is plainly an undeserved gift. If wisdom, righteousness, and life from God, for instance, are undeserved gifts to humans, then humans are in no position to take credit for these gifts. Misguided human pride

regarding such gifts is harmful, because it disallows God, from a specific human perspective, to be truly gracious in some central areas of human existence. Substituting human self-credit for divine grace, such pride obscures the importance of human dependence on a gracious, morally perfect God.

Because human agency is a requirement for divine companionship with humans, we should not expect divine grace to coerce a human response of commitment to life with God. Owing to moral weakness, however, humans cannot live up to God's moral character by themselves. They therefore fall short of perfectly obeying divine commands to love God fully and to love others unselfishly. In manifesting this human moral inadequacy, divine love commands show the human need for divine grace. They show that humans do not merit, or earn, approval from God, if God's moral character sets the standard for approval.

Human moral pride will not welcome this lesson, but this is no count against the lesson. Instead, such pride itself falls under proper suspicion once the reality of human moral deficiency emerges. This deficiency does not entail despair, however, if divine mercy underwrites forgiveness of humans.

Divine moral perfection combined with human moral imperfection, including human self-pride and despair, would call for a severe human struggle in receiving divine grace and wisdom. Following Pascal (*Penseés* sec. 435), we shall see that neither human self-pride nor human despair would enjoy a firm footing relative to a severe God's moral character. If divine grace and wisdom, as personified in Christ crucified, encompass redemptive self-sacrifice for God's purpose, then human appropriation of such grace and wisdom will be not only a response to, but also an exemplification of, such sacrifice. This appropriation will go against any human self-pride or despair incompatible with such sacrifice. It will involve, as a fitting response, self-sacrificial appropriation of grace and wisdom which are likewise self-sacrificial.

Gethsemane method

We can use expectation-evoking questions to focus on human expectations for God's reality, while offering the term "God" as a preeminent title requiring worthiness of worship. Our understanding

of evidence, and thus of knowledge, regarding God should be guided by the relevant notion of a God worthy of worship, rather than by a cognitive standard that is questionable relative to this notion.

Our initial question becomes not so much whether God exists as what the character and purposes of a God worthy of worship would be, if God exists. In keeping with the previous expectation-evoking questions, we should ask this: what *kind* of evidence and knowledge of God's reality should we expect a God worthy of worship to offer to humans?

A plausible answer is that God would offer the kind of evidence and knowledge that represents and advances God's worship-worthy moral character among humans.

We should expect evidence and knowledge of divine reality to be available to humans only in a manner suitable to divine *purposes* in self-revelation. We should expect these divine purposes to include the transformation of human moral characters toward God's moral character, for the sake of human improvement in companionship with God.

This lesson yields a major shift in our understanding of human knowledge and evidence of divine reality. It demands that inquirers become sensitive to the character and purposes of a God worthy of worship, in a manner that challenges and reorients human wills, as in Gethsemane. This lesson illuminates some of the severity in God and in human life, in terms of the needed transformation of humans in companionship with God.

Some philosophers think of religious faith as purely intellectual, similar to belief that, for instance, transfinite cardinal numbers or subatomic leptons exist. This, however, is a big mistake that leads to confusions about what kind of evidence is to be expected for religious faith. We shall see that the relevant evidence is to be appropriated in a self-sacrificial struggle in response to a divine challenge. Religious faith is a human moral struggle in appropriation of divine grace and wisdom, for the sake of reverent companionship with God.

Our inquiry about divine severity does not reduce to an inquiry about a traditional problem of evil, because the reality of divine severity is no challenge to God's existence or even God's goodness. On the contrary,

one should expect divine severity upon reflection about the implications of the worthiness of worship suited to God. Even so, one might propose that divine severity is a challenge to God's graciousness, on the ground that it suggests that God is holding back on the best for humans even if God is morally good on balance. A natural question, then, is this: could not God be more gracious and less severe, without any loss of value? In other words, could not God's "good and perfect will" (Romans 12, 2) be less severe? More specifically, is God at best restrictedly gracious toward at least some people? (Such questions occupy Moser 2013).

Divine severity, properly understood, does not undermine perfect divine grace but should be expected in the light of such grace. It points us to the volitional crisis of Gethsemane, for the sake of cooperative and lasting human life with God. In doing so, it invites us to consider the priority of divine power over philosophical propositions, persons over explanations, and God's will over human wills.

Accordingly, we should reconceive the philosophy of religion in the light of the Gethsemane crisis, including in the significant areas of the methodology and epistemology of God. This reconceiving leaves us with philosophy of religion renewed by a needed interpersonal and existential vitality, grounded in widely neglected but nonetheless salient evidence of God's redemptive severity.

If God is redemptively severe in a way that challenges human wills, then an important autobiographical question emerges: am I present for God's redemptive presence?

That is, am I genuinely available to struggle to receive and then to manifest God's purportedly transformative *agapē*? Or, alternatively, am I preoccupied with other priorities that omit divine *agapē* and a struggle for it as a priority in my life? If God's presence involves a kind of powerful *agapē* that includes an authoritative call to receive and obey God, am I willing to be adequately attentive to listen for and receive this call? In particular, is my conscience sufficiently sensitive and receptive to this life-giving call, particularly its offer and demand of unselfish *agapē*, in such a way that it enables me to hear and obey God's morally profound call? (On a divine call and the combination of offer/gift and demand regarding divine *agapē*, see Brunner 1937 (114-31 198-207); 1950 (183-99).

If I am sincerely receptive, I will be available to join the *agapē* struggle in a way that resists, as in Gethsemane, the priorities contrary to this divinely commanded struggle. To join this struggle is to join God's own life of ongoing struggle for self-giving love for all agents. Arguably, Jesus had this kind of divine expectation in mind with his striking remark that "there is need of only one thing" (Luke 10, 42).

We gain nothing by begging the question about God's existence. Instead, the present recommendation is to be both sincerely open to a subtle divine invitation to *agapē* struggle and genuinely attentive to any indicators of divine involvement for the sake of such struggle.

A way of testing the redemptive theism in question is to accept that recommendation with honesty, and then to assess one's subsequent experience accordingly over time. In this case, God ultimately must deliver the needed experiential evidential at the suitable time for each person. If the best explanation of one's experience supports such theism, then one can move beyond fideism and question begging (see Moser 2010 chaps. 2-4); (cf. Wiebe chap. 5).

Given that God's character of *agapē* is not purely conceptual, the needed evidential support for humans must have a basis in human experience. The previous considerations raise the issue of how we might remove obstacles between God and us, to participate in God's powerful life of *agapē* for others, including God's redemptive suffering for the world. Such questions also prompt the question of whether God's suffering for the world, such as in the self-giving death of Jesus for others (Romans 5, 8); (cf. Moser 2008 chap. 3) (2010 chap. 4), aims not to preserve or secure the world *as it is* but rather to redeem it in a new mode of life where *agapē* for others, including enemy-love, is the norm and the operative power for all agents.

We are now on conceptual soil foreign to traditional philosophy of religion, as we should expect in the presence of a challenging God worthy of worship³.

The primary concern of a truly redemptive God would be not so much with the specific intellectual content embraced by humans as with their

³ On the latter point, see (1 Corinthians 1, 18-25); (cf. Hays chap. 1) (cf. Gorman 275-80).

willingly joining an *agapē*-deepening struggle for others, in dependence on the God who is the source of perfect *agapē* for humans.

This struggle would be as much a struggle (in cooperation) *with* God as a struggle *for* (the redemptive mission of) God. In the divine remedial school, God would struggle with humans for the sake of their *agapē*-oriented transformation, both when they resist (at least without final rejection) and when they cooperate (if imperfectly). In other words, God would participate in the *agapē* struggle for the sake of disseminating and deepening the noncoercive power of divine *agapē* among humans. To that end, God would engage in self-giving suffering for the benefit of humans, as important parts of Jewish and Christian theology suggest (on which see Heschel (1962) Fiddes (1988)).

Given *agapē*-focused redemptive purposes for humans, God would aim to have us be in a distinctive relation to truth, in order to accommodate the *reality* of the purported redemption anchored in divine *agapē*. The aim would be, not just that we know truth, but more importantly that we ourselves *become true* to (that is, in full agreement with) divine *agapē*, because the intended redemption in *agapē* is for *us ourselves*, and not just for our beliefs. Kierkegaard (1991) puts the point as follows:

[...] Christianly understood, truth is obviously not [just] to know the truth but to be the truth... [O]nly then do I in truth know the truth, when it becomes a life in me. (205-6)⁴

The reality of divine *agapē*, then, must become a life, not just an object of reflection, in us. Divine severity can serve that redemptive end, as long as humans cooperate.

The suggested incarnational approach to redemptive truth goes beyond any merely intellectual relation to truth. It requires that the reality of divine *agapē* become the motivational center of our lives in order to have us become personally reflecting, or imaging, of God's moral character, ever more deeply.

This incarnational relation is *personifying* toward redemptive truth, and not merely intellectual, because divine *agapē* exceeds intellectual matters and involves a human *as a personal agent* with a life to live in the face of severity. The purported redemption in *agapē* involves cooperative

⁴ See Moser and McCreary (2010).

personal transformation toward God's moral character, and not just new knowledge, in virtue of a person's receiving and manifesting divine *agapē* for all persons, even one's enemies. The model, again, is Gethsemane.

The incarnational approach portrays God as coming to humans by judging, or subjecting to futility, all that is anti-God. The divine aim is to make willing humans new in the moral image and the companionship of the God who puts *agapē* first, even toward enemies.

This redemptive judgment works via willing human appropriation of redemptive truth whereby one struggles to participate in God's moral character and life, including in divine unselfish *agapē*. The litmus test for the divine authenticity of this struggle available to humans is the *experienced realization* of the aforementioned *agapē* struggle, including a struggle for enemy-love. In the absence of a struggle for enemy-love, in particular, one is not engaged in a struggle for or with the God worthy of worship. Accordingly, much of what goes under the category of "religion" does not positively involve the God worthy of worship. Much of it is, in fact, counterfeit and anti-God, despite any appearance or language of piety (see Juergensmeyer (2003) Stern (2003)); (cf. Seibert 2009).

When one is willing, the *agapē* struggle takes one beyond mere wishful thinking about God to the Gethsemane transformation of humans in volitional companionship with God. This struggle involves a meeting with God that includes an invitation and a demand: an invitation to life with (including in cooperation with) God and a demand to be conformed to God's moral character of unselfish *agapē*. Some people acknowledge a call to such a divine-human meeting; others do not. The resulting disagreement is familiar and undeniable, and it should come as no surprise on reflection.

The failure of some humans to acknowledge God's presence in the *agapē* struggle arises from various sources; as a result, we have no simple explanation on this front. Many different voices populate this world, and the remarkable multiplicity leaves many people confused and doubtful.

The cacophony leaves people not knowing what the needed redemptive gift of God would sound or look like. In addition, many people opt for selfishness and for hate of enemies over genuine *agapē*, and this choice creates a bias against acknowledgment of a God committed to *agapē*

toward all agents. Some people even announce their desire not to be in a world governed by God; out of fear of a loss of their autonomy (Nagel 130).

We should distinguish, however, between autonomy as independence of God and autonomy as capability to choose for or against a life of *agapē*. The latter autonomy is genuinely good and would be preserved by a God worthy of worship, but the former autonomy is arguably not good at all, given that humans lack the power to sustain their lasting flourishing on their own.

We must pay careful attention to the whole range of our experience, because the indication of God's reality should be expected to be subtle and even elusive. We should not expect divine revelation to be cheap and easy, given its unsurpassed value and profundity and our need for transformation toward divine *agapē*. A key issue is whether we can find self-giving suffering love at work, because divine *agapē* toward wayward humans would be inherently marked by it. This would be intentional *agapē* for the sake of others, grounded in one's looking to practice *agapē* for the good of others even when severity and suffering come to oneself.

What ultimately matters for the divine redemption of humans is *God's power of agapē*, not any human metaphysics, epistemology, philosophy, hermeneutics, theology, or other kind of human theory. We have considered the view, widely neglected among philosophers and others, that God's redemptive power is available to humans in the struggle of obedient receptivity toward the offer and demand of unselfish *agapē*, and not in mere thinking, talking, theorizing, or even arguing.

This is *dynamic, Gethsemane theism*, because it identifies God's redemptive power and corresponding self-revelation and evidence in connection with the struggle of human receptivity and activity, including active obedience on the part of humans. It assumes that God's redemptive power must give us humans, without coercion, the life we need if we are to survive and flourish lastingly in the face of severity.

According to Gethsemane theism, our lasting well-being comes from someone other than a mere human, as a humanly unmerited gift given noncoercively to willing humans at God's appointed time. This gift, however, comes with judgment on the world that is anti-God, not judgment as condemnation, but judgment as subjecting to futility all

that is anti-God in order to bring flourishing to willing humans in companionship with God. In coming with such redemptive judgment, this divine gift must be received in the aforementioned *agapē* struggle if it is to be received at all. The redemption on offer is thus dynamic rather than static.

One's being evidentially assured of God's reality and role in the *agapē* struggle is diachronic rather than once and for all, or synchronic, given that the divinely desired character transformation occurs over time, and not all at once. Even though the beginning of the struggle occurs at a particular moment, the decisive evidence in the struggle builds over time as a person struggles willingly and increasingly deeply with and for God.

Accordingly, the *agapē* struggle is an interactive person-to-person relationship that develops over time. Its characteristic evidence likewise yields knowledge of God over time, contrary to the familiar philosophical norm of knowledge of God as synchronic.

We cognitively limited humans do not have, and should not expect to have, a full explanation of the world's tragic and horrifying evil, but we still can have a perfectly loving sustainer with us in this world. Paul puts the point as the fact that nothing is able to separate willing humans from the *agapē* of God, "neither tribulation, distress, persecution, famine, peril, death, nor any other thing" (Romans 8, 35). Such a "theodicy" of *nonseparation* from God's *agapē* is ultimately the only theodicy on offer for us.

We now can begin to understand the otherwise cryptic remark at the center of the ministry of Jesus: "*Struggle* to enter through the narrow door, because many, I tell you, will seek to enter and not be able" (Lukas 13, 24).

Famously, Socrates remarked that the unexamined life is not worth living, but we now can add that the *nonstruggling* life relative to divine *agapē* is not worth living either. Philosophy of religion, then, should make room for the kind of wisdom that includes a divinely offered and commanded *agapē* struggle and its corresponding distinctive epistemology.

This sea change would yield profound benefits for the philosophy of religion as we know it. The test of authenticity, ultimately, is in the

living through the *agapē* struggle on offer. The remaining question is whether we humans are sincerely willing to undergo the test. Philosophy of religion now becomes existential and self-engaging, even urgent to human life.

BIBLIOGRAPHICAL REFERENCES

Bonhoeffer, Dietrich. *Discipleship: Dietrich bonhoeffer works, Vol. 4*. Minneapolis: Fortress Press, 2001. Print.

Brunner, Emil. *The divine imperative*. London: Lutterworth Press, 1937. Print.

---. *The Christian doctrine of God*. London: Lutterworth Press, 1950. Print.

Fiddes, Paul. *The creative suffering of God*. Oxford: Oxford University Press, 1988. Print.

Gorman, Michael J. *Cruciformity*. Grand Rapids: Eerdmans, 2001. Print.

Hays, Richard B. *First Corinthians*. Louisville: John Knox, 1997. Print.

Heschel, Abraham. *The prophets*. New York: Jewish Publication Society, 1962. Print.

Juergensmeyer, Mark. *Terror in the mind of God*. Berkeley: University of California Press, 2003. Print.

Kierkegaard, Søren. *Practice in christianity: Kierkegaard's writings, XX*. Princeton: Princeton University Press, 1991. Print.

Moser, Paul K. *The elusive God: Reorienting religious epistemology*. Cambridge: Cambridge University Press, 2008. Print.

---. *The evidence for God: Religious knowledge reexamined*. Cambridge: Cambridge University Press, 2010. Print.

---. *The severity of God: Religion and philosophy reconceived*. Cambridge: Cambridge University Press, 2013. Forthcoming.

Moser, Paul K. and Mark McCreary. "Kierkegaard's conception of God". *Philosophy compass*. 2010: 127-35. Print.

Nagel, Thomas. *The last word*. New York: Oxford University Press, 1997. Print.

Pascal, Blaise. *Pensées*. London: Penguin, 1995. Print.

Paul K. Moser

Seibert, Eric A. *Disturbing divine behavior: Troubling Old Testament images of God*. Minneapolis: Fortress Press, 2009. Print.

Stern, Jessica. *Terror in the name of God*. New York: HarperCollins, 2003. Print.

Wiebe, Phillip. *God and other spirits*. New York: Oxford University Press, 2004. Print.

Wittgenstein, Ludwig. *Notebooks 1914-1916*. Chicago: University of Chicago Press, 1984. Print.



EPISTEMIC VS ONTIC CLASSIFICATION OF QUANTUM ENTANGLED STATES?

¿LA CLASIFICACIÓN ONTICA VS CLASIFICACIÓN EPISTÉMICA,
DE LOS ESTADOS CUÁNTICOS ENTRELAZADOS?

MICHELE CAPONIGRO

Indeterminism in Sciences and Historico-philosophical Transdisciplinary
Advanced Research Center, Italia.

Universidad de Bergamo, Italia. michelecaponigro@unibg.it

ENRICO GIANNETTO

Universidad de Bergamo, Italia. enricogiannetto@unibg.it,

RECIBIDO EL 20 DE FEBRERO DE 2012 Y APROBADO EL 6 DE ABRIL DE 2012

RESUMEN

En este breve artículo, y a partir de trabajos recientes, se analiza desde el punto de vista conceptual la siguiente pregunta básica: ¿puede la naturaleza de los estados cuánticos entrelazados ser interpretada ontológica o epistemológicamente? De acuerdo a algunos trabajos especializados, los grados de libertad (y la herramienta de la partición) de los sistemas cuánticos, nos permite establecer una posible clasificación entre los estados entrelazados y los estados factorizables. Sugerimos que la "elección" de los grados de libertad (o partición cuántica), incluso si se justifican matemáticamente, introducen un elemento epistémico, no sólo en los sistemas, sino también en su clasificación. Planteamos, en cambio, que no existen dos clases de estados cuánticos: entrelazados y factorizables sino tan solo una clase de estados: los entrelazados. De hecho, los estados factorizables se vuelven entrelazados para una elección diferente de sus grados de libertad (es decir, que están entrelazados con respecto a otros observables). De la misma manera, no hay particiones de sistemas cuánticos que posean un status ontológicamente superior con respecto a otros. Por todas estas razones, ambas herramientas matemáticas utilizadas (i.e., las particiones cuánticas y los grados de libertad) son responsables de la creación de una clasificación incorrecta de los sistemas cuánticos. Finalmente, sostenemos que no se puede hablar de una clasificación de los sistemas cuánticos: todos los estados cuánticos muestran una naturaleza objetiva única, son todos estados entrelazados.

PALABRAS CLAVE

Elementos epistémicos vs elementos ónticos, entrelazamiento cuántico, subsistemas (particiones y estados únicamente factorizables).

ABSTRACT

In this brief paper, starting from recent works, we analyze from a conceptual point of view this basic question: can the nature of quantum entangled states be interpreted ontologically or epistemologically? According to some works, the degrees of freedom (and the tool of quantum partitions) of quantum systems permit us to establish a possible classification between factorizable and entangled states. We suggest, that the "choice" of degree of freedom (or quantum partitions), even if mathematically justified introduces an epistemic element, not only in the systems but also in their classification. We retain, instead, that there are not two classes of quantum states, entangled and factorizable, but only a single class of states: the entangled states. In fact, the factorizable states become entangled for a different choice of their degrees of freedom (i.e. they are entangled with respect to other observables). In the same way, there are no partitions of quantum systems which have an ontologically superior status with respect to any other. For all these reasons, both mathematical tools utilized (i.e. quantum partitions or degrees of freedom) are responsible for creating an improper classification of quantum systems. Finally, we argue that we cannot speak about a classification of quantum systems: all quantum states exhibit a uniquely objective nature, they are all entangled states.

KEY WORDS

Epistemic vs. ontic elements, quantum entanglement, subsystems (partitions and uniquely factorizable states).

I Systems and partitions

In spite of continuous progress, the current state of entanglement theory is still marked by a number of outstanding unresolved problems. These problems range from the complete classification of mixed-state bipartite entanglement to entanglement in systems with continuous degrees of freedom, and the classification and quantification of multipartite entanglement for arbitrary quantum states.

In this paper, starting from two important works, i) Torre (2010) and ii) Zanardi (2001), we will analyze the possible relationship among these elements:

- 1) The degrees of freedom of the quantum system.
- 2) The partitions of the quantum system.
- 3) The epistemic elements introduced from the procedures (1) and (2).

As we know, the relationship between quantum systems (QS) and their possible quantum entangled systems (QES) is not a trivial question. There are many efforts to understand this dynamics. Zanardi (2001) in his paper argues that the partitions of a possible system do not have an ontologically superior status with respect to any other: according to Zanardi given a physical system S , the way to subdivide it in subsystems is in general by no means unique. We will analyze his conclusion in the following sections below.

According to Zanardi the consequences of the “non uniqueness of the decomposition of a given system S into subsystems” imply (at the quantum level), a fundamental ambiguity about the very notion of entanglement that accordingly becomes a relative one. The concept of “relative” for an entangled system has been developed by Viola and Barnum (2006).

They concentrate their efforts on this fundamental question: how can entanglement be understood in an arbitrary physical system, subject to arbitrary constraints on the possible operations one may perform for describing, manipulating, and observing its states? In their papers, the authors proposed that entanglement is an inherently “relative concept”, whose essential features may be captured in general in terms of the relationships between different observers (i.e. expectations of quantum observables in different, physically relevant sets).

They stressed how the role of the “*observer*” must be properly acknowledged in determining the distinction between entangled and unentangled states.

Quantum entanglement: brief overview

From a phenomenological point of view, the phenomenon of entanglement is quite simple. When two or more physical systems form an interaction, some correlation of a quantum nature is generated between the two of them, which persists even when the interaction is switched off and the two systems are spatially separated.

Quantum entanglement describes a non-separable state of two or more quantum objects and has certain properties which contradict common physical sense. While the concept of entanglement between two quantum systems, which was introduced by Erwin Schrödinger (1936), is well understood, its generation and analysis still represent a substantial challenge. Moreover, the problem of quantification of entangled states; is a long standing issue debated in quantum information theory. Today the bipartite entanglement (“two-level systems”, i.e. *qubits*) is well understood and has been prepared in many different physical systems. The mathematical definition of entanglement varies depending on whether we consider only pure states or a general set of mixed states (Giannetto)¹. In the case of pure states, we say that a given state $|\psi\rangle$ of n parties is *entangled* if it is not a tensor product of individual states for each one of the parties, that is,

$$|\psi\rangle \neq |v_1\rangle_1 \otimes |v_2\rangle_2 \otimes \dots \otimes |v_n\rangle_n \quad (1)$$

For instance, in the case of 2 qubits A and B (sometimes called “Alice” and “Bob”) the quantum state

$$|\psi^+\rangle = \frac{1}{\sqrt{2}} [(|0\rangle_A \otimes |0\rangle_B + |1\rangle_A \otimes |1\rangle_B)] \quad (2)$$

¹ Where it is discussed the reason why entanglement generally requires a density matrix formalism.

is entangled since $|\psi^+\rangle \neq |v_A\rangle_A \otimes |v_B\rangle_B$. On the contrary, the state

$$|\Phi\rangle = \frac{1}{2} [(|0\rangle_A \otimes |0\rangle_B + |1\rangle_A \otimes |0\rangle_B) + |0\rangle_A \otimes |1\rangle_B + |1\rangle_A \otimes |1\rangle_B] \quad (3)$$

is not entangled, since

$$|\Phi\rangle = \left(\frac{1}{\sqrt{2}}(|0\rangle_A + |1\rangle_A)\right) \otimes \left(\frac{1}{\sqrt{2}}(|0\rangle_B + |1\rangle_B)\right) \quad (4)$$

A pure state like the one from Eq. 2 is called a *maximally entangled state of two qubits*, or a *Bell pair*, whereas a pure state like the one from Eq. 4 is called *separable*. In the general case of mixed states, we say that a given state p of n constituent states is *entangled* if it is not a probabilistic sum of tensor products of individual states for each one of the subconstituents, that is,

$$P \neq \sum_k p_k p_1^k \otimes p_2^k \otimes \dots \otimes p_{F_c}^k, \quad (5)$$

with $\{p_k\}$ being some probability distribution. Other-wise, the mixed state is called *separable*. The essence of the above definition of entanglement relies on the fact that entangled states of n constituents cannot be prepared by acting locally on each one of them, together with classical communication among them. Entanglement is a genuinely quantum-mechanical feature which does not exist in the classical world. It carries non-local correlations between the different systems in such a way that they cannot be described classically.

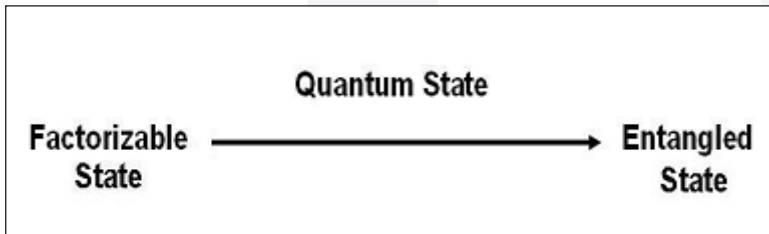


Figure 1. Torre's main thesis: factorizable states become entangled in different degrees of freedom.

II

Are quantum states all entangled?

As mentioned above, the recent work by Torre (2010) is a fundamental paper which gives us the possibility to speculate about the nature and the classification of entangled states. The paper demonstrates that a state is factorizable in the Hilbert space corresponding to some choice of degrees of freedom, and that this same state becomes entangled for a different choice of degrees of freedom.

Therefore, entanglement is not a special case, but is ubiquitous in quantum systems. According to the authors, one may erroneously think that there are two classes of states for the QS: 1) factorizable and 2) entangled, which correspond to qualitative difference in the behavior of the system, close to classical in one case and with strong quantum correlations in the other. They argue that this is indeed wrong because factorizable states also exhibit entanglement with respect to *other* observables. In this sense, all states are entangled; “entanglement is not an exceptional feature of some states but is ubiquitous in QM”.

To sum up this conceptual analysis by Torre and Zanardi, we think that there is an unclear relationship among these elements:

1. Factorizable states (Torre).
2. Entangled states.
3. The (choice) of partitions of quantum system (Zanardi).
4. The role of the observer (in determining the distinction between entangled and unentangled states).

We think that all points (except the second point) introduce epistemic elements in the analysis and in the classification of the quantum systems. We suggest that the second point is the key to understand the nature of the underlying physical reality. We argue in the next sections, that the conceptual analysis of Torre and Zanardi differs from what we suggest concerning the epistemic elements introduced in their papers.

Factorizability of a state as epistemic property?

An important question is related at the property of factorizability of quantum state. Is the factorizability tool an objective property? Briefly stated, is factorizability an objective property of the system or is it a feature of (our) description of system (i.e. an epistemic property?)

With reference to Torre's paper (2010), the authors show that factorizability and entanglement "are not preserved" in a change of the degrees of freedom used to describe the system, they demonstrate in details case that the factorizability of a state is a property that is "not" invariant under a change of the degrees of freedom that we use in order to describe the system. From mathematical point of view (1), they consider a quantum system with two subsystems $S = (S_A; S_B)$ that may correspond to two degrees of freedom A and B . The state of the system belongs to the Hilbert space $H = H_A \otimes H_B$ and the two degrees of freedom are represented by operators $A \otimes I$ and $I \otimes B$. Suppose it is given that the system has a factorizable, none entangled, state $\Psi = \Psi_A \otimes \Psi_B$ with Ψ_A and Ψ_B arbitrary states (not necessarily eigenvectors of A and B) in the spaces H_A and H_B . Then there exists a transformation of the degrees of freedom $F = F(A, B)$ and $G = G(A, B)$ that suggests a different factorization, $H = H_F \otimes H_G$, where the state is no longer factorizable: $\Psi \neq \Psi_F \otimes \Psi_G$ with $\Psi_F \in H_F$ and $\Psi_G \in H_G$. The state becomes entangled with respect to the new degrees of freedom; the factorizability of states is not invariant under a different factorization of the Hilbert space.

To conclude, they have shown that for any system in a factorizable state, it is possible to find different degrees of freedom that suggest a different factorization of the Hilbert space where the same state becomes entangled; for this reason they argued that every state, even for those factorizable, it is possible to find pairs of observables that will violate Bell's inequalities. The figure above (n. 1) summarize Torre's thesis.

The authors analyze also the inverse problem: the fact that the appearance of entanglement depends on the choice of degrees of freedom can find an interesting application in the "disentanglement" of a state; one can, sometimes, transform an entangled state into a factorizable one by a judicious choice of the degrees of freedom.

To conclude, we think that the epistemic element is inherent in the possibility to "choose" the degrees of freedom of the quantum system: this possibility affects the classification of quantum states in entangled or factorizables. In fact, it is simple to ask these epistemological questions: a) what are the degrees of freedom for a quantum system? b) Is it a complete set that describe all quantum properties? Can be a particle entangled in one context be factorizable in another context?

The partitions of quantum system as epistemic property?

As we have seen, given a quantum system, the way to subdivide (to partition) it in subsystems in “not unique”. We call this first phase “epistemic”, as in fact we are able to decide how to partition the quantum system.

The conclusion of this operation is most important of its premise: in fact if we find (in the subsystems) an entangled state, this state has an ontological nature “but only if referred” to that kind of particular partition. We have, in other words, an objective entangled state for an epistemic partition! For these reasons, the notions of an entangled state become a relative concept and the relativity of this concept is linked to us, at the choice of partitions or degrees of freedom. At the same time, the property of the entangled state is objective. The figures above (n. 2) represent our view of Zanardi’s problem.

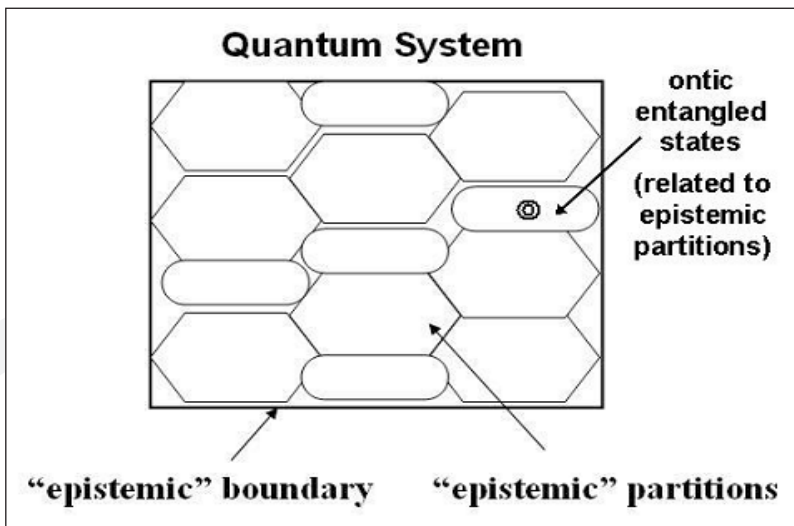


Figure 2. Epistemic partitions and ontic entanglement.

III

Some considerations and conclusions

We have seen that quantum systems admit a variety of tensor product structures depending on the complete system of commuting observables chosen for the analysis; as a consequence we have different notions of entanglement associated with these different tensor products.

We notice that, in the determination of whether a state is factorizable or entangled, the factorization of the Hilbert space is crucial and this factorization depends on the choice of the observables corresponding to the degrees of freedom. In the same way, as Zanardi stressed, given a quantum system, the way to subdivide it (via partitions) in subsystems it is “not unique”; the partitions of a possible system have not an ontologically superior status with respect to any other.

Based on these points, we argue that the criteria of partitions and factorizability (or partitions) contain an *a priori* epistemic element, the figure (n. 3) summarize our position.

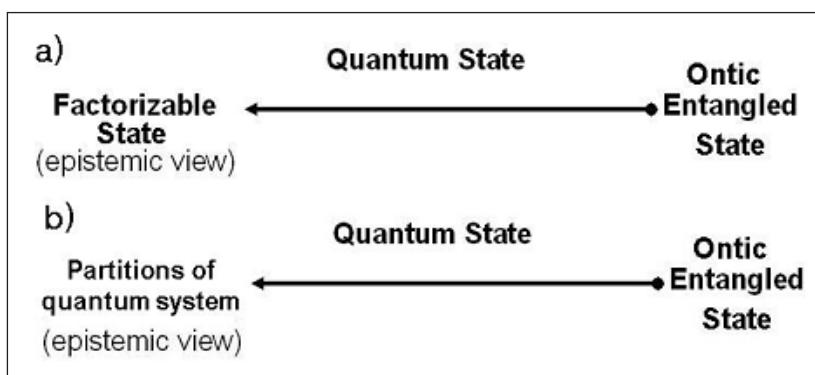


Figure 3. Our position.

In conclusion, we suggest that all quantum system exhibit an objective nature that is entangled, at basic level the underlying physical reality is entangled. A quantum state could be non-entangled if and only if it would be factorizable for every possible partition or choice of degrees of freedom, but this can never occur. The epistemic level emerges with the “observer” (partitions or degree of freedom) the physicists and philosophers should consider these arguments in their debates.

BIBLIOGRAPHICAL REFERENCES

De la Torre, Alberto Clemente et al. “Entanglement for all quantum states” *Eur. J. Phys.* 2010: 325-332. Print.

Giannetto, Enrico. “Some remarks on non-separability”. *The foundations of quantum mechanics: Historical analysis and open questions*. Claudio

Garola and Arcangelo Rossi (eds.). Dordrecht: World Scientific Pub Co Inc, 1995: 315-324. Print.

Schrödinger, E. "Discussion of probability relations between separated systems". *Proceedings of the Cambridge Philosophical Society*. Oct. 1936: 555-563. Print.

---. "Discussion of probability relations between separated systems". *Proceedings of the Cambridge Philosophical Society*. 1936: 446-451. Print.

Viola, Lorenza and Howard Barnun. "Invited contribution to the Proceedings of the Boston Colloquium for Philosophy of Science on "Foundations of Quantum Information and Entanglement". *Quantum Physics*. Boston: Cornell University Library, 2006. Online.

Zanardi, Paolo. "Virtual quantum subsystems". *Phys. Rev. Lett.* 2001. Print.

LA IDEALIZACIÓN EN LA MATEMÁTICA*

IDEALIZATION IN MATHEMATICS

THOMAS MORMANN

University of Basque Country, País Vasco. ylxmormot@ehu.es

RECIBIDO EL 30 DE MARZO DE 2012 Y APROBADO EL 17 DE MAYO DE 2012

RESUMEN ABSTRACT

El objetivo del presente artículo consiste en elucidar el papel de las idealizaciones en la evolución del conocimiento matemático, inspirado por algunas ideas de la filosofía de la ciencia y de la matemática neokantiana de Ernst Cassirer. Usualmente, en la filosofía de la ciencia contemporánea se da por hecho que el asunto de la idealización tiene que ver únicamente con las idealizaciones en las ciencias empíricas, en particular en la física.

Por contraste, Cassirer sostuvo que la idealización en las matemáticas, así como en las ciencias, tiene la misma base conceptual y epistemológica. Más precisamente, examino su "tesis de la identidad" investigando una variedad de ejemplos de idealizaciones tomadas del álgebra, la topología, la teoría de redes y la geometría física. Las idealizaciones en las matemáticas, así como en el conocimiento físico, se pueden caracterizar por la introducción de elementos ideales que conducen a compleciones. En ambas áreas, estos elementos ideales desempeñan esencialmente el mismo papel, es decir, sustituyen una variedad incompleta de objetos, mediante una variedad conceptual completa "idealizada".

The aim of this paper is to elucidate the role of idealizations in the evolution of mathematical knowledge inspired by some ideas of Ernst Cassirer's Neokantian philosophy of science and mathematics. Usually, in contemporary philosophy of science it is taken for granted that the issue of idealization is concerned only with idealizations in the empirical sciences, in particular in physics.

In contrast, Cassirer contended that idealization in mathematics as well as in the sciences has the same conceptual and epistemological basis. More precisely, this "sameness thesis" is scrutinized by investigating a variety of examples of idealizations taken from algebra, topology, lattice theory, and physical geometry. Idealizations in mathematical as well as in physical knowledge can be characterized by the introduction of ideal elements leading to completions. In both areas these ideal elements play essentially the same role, namely, to replace an incomplete manifold of objects by a complete "idealized" conceptual manifold.

PALABRAS CLAVE KEY WORDS

Algebras booleanas, Cassirer, compleciones, cortaduras de Dedekind, idealización, espacios de Stones.

Boolean algebras, Cassirer, completions, Dedekind cuts, idealization, Stones spaces.

* Agradecimientos al Ministerio Español de Ciencia e Innovación por financiar la presente investigación, la cual hace parte del proyecto FFI 2009-12882

El problema de la idealización en la filosofía de la ciencia

El papel de la idealización en las ciencias empíricas y en particular, en la física es un tema muy discutido en filosofía de la ciencia Norton (2012), Nowakowa & Nowak (2000), Weisberg (2007). Sin embargo, esta misma noción ha recibido una menor atención en discusiones filosóficas relacionadas con las matemáticas. Hoy en día, al tratar el tema de la idealización en las ciencias se supone implícitamente que *en* las matemáticas no hay lugar para la idealización; éstas, ya se encuentran, en el lado de lo ideal. Así, el problema del carácter idealizador del conocimiento científico, solamente se plantea en relación con el papel de la idealización en el terreno empírico.

El filósofo Leszek Nowak y sus colaboradores en las últimas décadas, establecieron una detallada clasificación de los distintos métodos de idealización (2000). “Evidentemente”, cabría decir que sólo les preocupa la idealización en las ciencias empíricas.

En este artículo, basándome en las ideas del filósofo neokantiano Ernst Cassirer, argumentaré, que también en las matemáticas, la idealización desempeña un papel clave y además la separación estricta de métodos de idealización en las ciencias empíricas y la matemática no es nada recomendable. Según Cassirer, la forma de tratar el problema de la idealización, adopta un punto de partida demasiado tardío. Para él, la idealización tiene un papel, tanto en la matemática, como en las ciencias empíricas, por lo que una teoría sobre el papel de la idealización en la ciencia, debe tener en cuenta, *tanto* a la primera, *como* a las segundas.

Si queremos entender el papel de la idealización, debemos estudiar cómo funciona en las matemáticas y en las demás ciencias. En segundo lugar, debo señalar que para Cassirer la filosofía de la ciencia no tiene como tarea darle fundamentación a las ciencias, en el sentido de proporcionarles el suelo firme del que no disponen por sí mismas. De lo que se trata, es de entender a la ciencia filosóficamente, explicar lo que en la práctica científica puede parecer a menudo confuso y poco claro. Por tal motivo, la tarea de la filosofía de la ciencia es la reconstrucción racional, más no la fundamentación metafísica.

Para la filosofía de la ciencia de Cassirer, el punto esencial de referencia es la física matemática: no la matemática como ciencia de objetos ideales, ni la física como ciencia puramente empírica. Esto nos puede bastar como

prueba para demostrar que Cassirer, no pretendía una fútil reducción de la física a la matemática, ni una mera identificación entre ambas.

Si se me permite expresar esa relación (entre ciencia y filosofía de la ciencia) de forma algo burda y paradójica, diría que la atención de la filosofía no debe dirigirse ni a la matemática ni a la física, sino específicamente a la relación¹ entre esos dos ámbitos. (Cassirer 48)

El problema de la idealización en la matemática

Es un lugar común afirmar que las matemáticas tratan con objetos “ideales” y no con objetos “reales”. Tomemos por ejemplo, la geometría como teoría matemática: los puntos, las líneas y las áreas en geometría, *no* son reales, sino ideales. De forma más general, los objetos y las relaciones de la matemática, no son reales, sino objetos y relaciones ideales. No obstante, este no es el modo en el que quisiera utilizar los conceptos “ideal”, “idealización” y nociones similares. Por ello, usaré el concepto de idealización en el sentido de idealización relativa, es decir, consideraré que las idealizaciones siempre son idealizaciones de algo y relativas a algo. Aunque esto puede parecer oscuro, por ello, trataré de clarificarlo.

Hubo un tiempo en el que la idealización era un tema candente en matemáticas: los *elementos ideales*, hicieron su aparición en la geometría; la teoría de los números; la teoría de retículos y en muchos otros campos de la matemática. Esto puede apreciarse, aún hoy en día, por el uso en la matemática contemporánea de expresiones como: “punto ideal”; “número ideal” o sencillamente; “ideales” de varios tipos como: “ideales primos”, “ideales redondos” y muchos otros más.

La filosofía de la matemática en pocas ocasiones, ha prestado atención al tema de la idealización. Una excepción a esto es el caso del filósofo neo-kantiano alemán Ernst Cassirer. Según Cassirer, las idealizaciones son de vital importancia para la matemática y las ciencias empíricas.

Este artículo, presenta algunos aspectos esenciales de la propuesta de Cassirer sobre el papel que desempeña la idealización en las matemáticas y las ciencias empíricas. Estos aspectos desde mi posición, podrían resultar de interés para la filosofía de la ciencia, incluso, hoy en día.

¹ Subrayado por parte del autor.

La filosofía de la ciencia de Cassirer, trató a la matemática y a las ciencias, en particular a la física, como un “continuo”. Esto no quiere decir, que considerase a las matemáticas, como si fuesen un tipo de ciencia *empírica*, tal y como eran entendidas por Mill y sus modernos seguidores naturalistas. Más bien, la perspectiva de Cassirer, estaba basada en las siguientes tesis “idealistas”:

(**TI**) *Idealización en la matemática y en las ciencias (“tesis de la identidad”)*: la idealización en la matemática, así como en las ciencias empíricas matematizadas, puede ser caracterizada mediante la introducción de “elementos ideales”. En ambas áreas, la introducción de elementos ideales, juega esencialmente, el mismo papel.

(**IC**) *Idealización como completación*: la idealización en matemáticas, es completación, esto es, la completación de un dominio dado, mediante la introducción de algunos “elementos ideales”.

La tesis de la identidad (**TI**), fue introducida por Cassirer en uno de sus primeros artículos titulado “Kant und die moderne Mathematik” (“Kant y la matemática moderna”) de 1907 -hasta ahora, este artículo no está completamente traducido al inglés o al español-, este importante trabajo, puede ser considerado como el precursor programático de su primer gran libro: *Sustancia y función* de 1910, una de las obras más conocida de este autor.

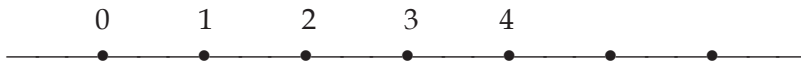
El objetivo general de este artículo, es mostrar que las completaciones idealizantes, son una realidad en la práctica matemática y no la invención de algunos filósofos. En pocas palabras, planteo que las completaciones, deben considerarse como herramientas esenciales en el arsenal de la “matemática de los matemáticos” y que merecen una mayor atención por parte de la filosofía de la matemática contemporánea.

De los números naturales N , a los números enteros Z : aspectos matemáticos y filosóficos

Comencemos con un ejemplo muy elemental, que sin embargo, jugó un papel importante, no sólo en la historia de las matemáticas, sino también, en la historia de la filosofía e incluso en la actualidad, puede ser considerado como un prototipo de una completación idealizadora.

Hoy en día, es difícil imaginar la dificultad conceptual que puede entrañar, pasar de los números naturales \mathbf{N} $\{0, 1, 2, 3, \dots\}$ a los números enteros \mathbf{Z} $\{\dots, -2, -1, 0, +1, +2, \dots\}$. Pero este no siempre ha sido el caso. Si consultamos la historia de las matemáticas, algunos filósofos y matemáticos, percibieron dificultades filosóficas incluso con el *zero* y el *uno*, pero dejemos de lado esta cuestión y tomemos el conjunto de los números naturales \mathbf{N} , como algo dado y no problemático, dotado con las operaciones estándares de la adición y la multiplicación.

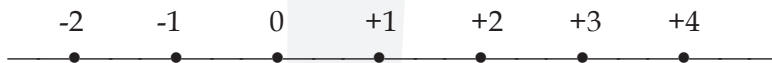
Un modelo geométrico útil de los números naturales \mathbf{N} , representa estos números como puntos equidistantes sobre una línea recta:



Este modelo nos muestra cierta asimetría de \mathbf{N} . Sólo los puntos, que están en el lado derecho de un punto *zero*, que antes hemos elegido de una manera arbitraria, sirven como representantes de los números naturales. Por tanto, es natural preguntarse: ¿qué se puede decir sobre los puntos que se encuentran a la izquierda de 0?

Dicho de una manera más formal, los números naturales se muestran estructuralmente incompletos en el sentido de que algunas ecuaciones lineales, tienen solución en \mathbf{N} , mientras que otras *no*. Por ejemplo: $4x + 3y = 20$, tiene solución en \mathbf{N} , mientras que $4x + 3y = 2$, no tiene una solución en \mathbf{N} , aunque ambas ecuaciones parezcan esencialmente del mismo tipo. Como sabemos actualmente, los números enteros \mathbf{Z} , ofrecen una solución altamente satisfactoria a dicha problemática.

Para remediar estas deficiencias ya mencionadas, parece oportuno completar los números naturales \mathbf{N} con una estructura más simétrica, por ejemplo, es más completa la estructura de los números enteros \mathbf{Z} , la cual puede ser representada geoméricamente de la siguiente manera:



Remplazando el dominio \mathbf{N} por el dominio \mathbf{Z} , aparece una estructura, donde la ecuación $4x + 3y = 2$ sí tiene una solución, además de muchas otras ecuaciones que pueden resolverse y que no tienen solución en el

ámbito más pequeño de \mathbf{N} . Como es bien sabido, esta completación de \mathbf{N} a \mathbf{Z} , no es un mero juego formal, sino que posee aplicaciones útiles, por ejemplo: se pueden interpretar estos números negativos como deudas y los positivos como capital o tal vez, como la magnitud de una fuerza, en relación con otra fuerza opuesta, además de muchas otras maneras, como ya lo hemos aprendido en la escuela.

Aunque la completación de \mathbf{N} a \mathbf{Z} , hoy en día es considerada como una pequeña pieza de la matemática elemental, hace un tiempo atrás, era vista como un tema central para la filosofía. De hecho, puede ser considerada, como un paso decisivo al separar el idealismo trascendental emergente de Kant, del racionalismo de la escuela de Leibniz y Wolff.

A diferencia del racionalismo de Leibniz, Kant sostuvo que el ámbito de nuestra experiencia no debe entenderse como una versión más o menos confusa, del ámbito de lo racional. Para Kant, las esferas de la razón y la experiencia, son esencialmente diferentes y esa diferencia, se expresa mediante la idea de magnitud negativa (cf. Beiser).

Kant argumentaba que la esfera de la experiencia, no se puede describir en términos estrictamente lógicos. Prueba de esto, es que una importante característica del ámbito de la experiencia es la “oposición real” de las cosas, la cual es esencialmente distinta de la “oposición lógica” reinante en el ámbito de la razón. La idea de “oposición real” de Kant, estaba basada en el concepto matemático de *magnitudes negativas*.

El texto pre-crítico, donde Kant formula estas ideas, data de 1763 y se titula: “Versuch den Begriff der negativen Größen in die Weltweisheit einzuführen” (“Intento de introducir el concepto de las magnitudes negativas en la filosofía de la ciencia”). En este artículo, Kant habla sobre la necesidad de la completación de los números naturales \mathbf{N} , a los enteros \mathbf{Z} . El elemento central en su argumentación es la distinción entre *oposición lógica* y *oposición real*. La *oposición lógica*, es una simple *contradicción*, esto es, la relación entre un *predicado A* y su *negación no-A*. Así, la oposición lógica nos lleva a algo *imposible*, nos lleva a *afirmar* y a *negar* al mismo tiempo, el predicado *A*. La *oposición real*, debe ser estrictamente distinguida de la oposición lógica. Como explicó Kant:

Supongamos que hay +8 unidades de capital y -8 de la deuda pasiva, no hay contradicción involucrada en atribuirles a la misma persona. Sin embargo, una de estas

magnitudes cancela una cantidad que es igual a la postulada por el otro, y la consecuencia es cero. (Kant 174)

Lo que Kant quería resaltar en contra de los seguidores de Leibniz, era que relaciones, tales como: la relación entre *capital* y *deuda*, es decir, relaciones que apelan a una oposición real, *no* son contradicciones, debido a que estas, pueden suceder en la realidad. Por tanto, hay que distinguirlas de las relaciones que conllevan una oposición lógica entre el predicado *A* y su negación *no-A*, las cuales, *no* pueden suceder en la realidad.

Kant sostuvo, que la *oposición real* (*Realrepugnantz*), era un aspecto esencial de la realidad fenoménica. Así, la incapacidad del racionalismo de Leibniz-Wolff para captar adecuadamente la idea de *oposición no-lógica* era una deficiencia fatal. En comparación, su idealismo trascendental, no sufría dicha deficiencia, porque permitía abordar, tanto la oposición lógica, como la oposición real.

No obstante, esta versatilidad tenía un costo. La lógica (en el sentido tradicional), carece del poder conceptual necesario para controlar la realidad en la cual la oposición real, juega un papel esencial.

Para tratar oposiciones reales, parece ser necesario acudir a otros recursos no-lógicos en concreto, cierta forma de geometría rudimentaria que incluye la distinción entre “izquierdo” y “derecho”, tal y como se muestra en el modelo geométrico de los números naturales y enteros, con el que hemos comenzado este acápite:



Entonces, el “8 derecho” (+8) y el “8 izquierdo” (-8) se cancelan, i.e. $(+8) + (-8) = (-8) + (+8) = 0$. Kant en su filosofía madura, identificó estos ingredientes no-lógicos, como las formas de la sensibilidad pura, en concreto como el espacio y el tiempo (Beiser 41-2).

La completación geométrica de los números naturales \mathbf{N} , mencionada hace unos momentos, seguramente era poco exacta. Por tanto, ahora volveremos una vez más, al concepto de completación, retomando el

ejemplo de la completación de los números naturales a los enteros y elaborando sus rasgos matemáticos, con más detalle.

Sería un error entender el carácter elemental de esta completación de los números naturales a los números enteros, como una prueba de su trivialidad. Voy a argumentar, que incluso en la matemática contemporánea, la cuestión acerca de este tipo de completaciones, tiene cierto interés. De hecho, la completación de los números naturales a los enteros, es sólo el caso más elemental, de lo que hoy se suele llamar “completación de Grothendieck” de un semigrupo conmutativo $(S, +)^2$.

Un semigrupo conmutativo, es un dominio S , dotado con una adición $+$, que más o menos se comporta como la adición de números naturales. Cabe destacar que un elemento a de S , no tiene necesariamente un inverso $-a$. La completación de dicho semigrupo, se puede esbozar de la siguiente manera. Dado un semigrupo $(S, +)$, el producto cartesiano $S \times S$, se dota canónicamente, con una *estructura de semigrupo* definida por:

$$(a, b) + (a', b') := (a + a', b + b') \text{ con } (a, b), (a', b') \in S \times S$$

En un segundo paso, definimos en $S \times S$, la siguiente relación de equivalencia

$$(a, b) \sim (a', b') := a + b' + s = b + a' + s \text{ para algún } s \in S$$

Denotemos $[a, b] := \{(a', b'); (a', b') \sim (a, b)\}$. Esta relación de equivalencia es compatible con la estructura del semigrupo de $S \times S$, es decir, también $S \times S / \sim$ es un semi-grupo:

$$[a, b] + [a', b'] = [a + b, a' + b']$$

La siguiente proposición, nos muestra que $S \times S / \sim$ es la completación deseada del semigrupo $(S, +)$:

Proposición: sea $(S, +)$, un semigrupo conmutativo con un elemento neutral O . Con base en esto, existe una inclusión canónica $S \xrightarrow{c} S \times S / \sim$ definida como $c(a) := [a, O]$, que preserva la estructura del semigrupo. La estructura del semigrupo, de $S \times S / \sim$, no sólo es una

² En 1872, Dedekind, conoció esta completación en el caso de los números naturales \mathbb{N} , Sieg & Schlimm (2005 136).

estructura de semigrupo, sino una estructura de grupo, esto es, cada elemento $[a, b]$, tiene su inverso $[b, a]$, tal que:

$$[a, b] + [b, a] = : [a + b, a + b] = [0, 0] \blacklozenge$$

Como cabe esperar, la completación de \mathbf{N} a \mathbf{Z} , esbozada informalmente, al principio de este texto es isomórfica a la completación de Grothendieck, de $\mathbf{N} \times \mathbf{N}/\sim$.

Además, la completación de Grothendieck, puede mostrarse como la “mejor” completación posible en el sentido de la teoría de las categorías, porque está basada en una construcción universal.

Resumiendo, la completación de \mathbf{N} a \mathbf{Z} , sólo es un ejemplo elemental de un proceso de completación conceptual más general, que es importante en muchos ámbitos de las matemáticas contemporáneas. Por ejemplo, la completación de Grothendieck, desempeña un papel crucial para la construcción de la “K-teoría” que resultó ser una herramienta significativa en la topología algebraica moderna y en el análisis funcional. En el caso de la K-teoría topológica, el semigrupo que se quiere completar es el semigrupo $V(X)$ de los haces vectoriales sobre un espacio topológico X . Por ende, se puede mostrar que la completación $K(X)$ de $V(X)$, figura como una componente de la teoría cohomológica periódica $K^*(X)$.

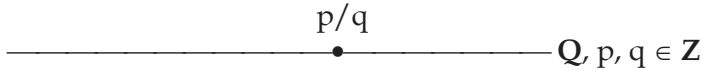
Las cortaduras de Dedekind

La completación de Grothendieck de los números naturales es un procedimiento de completación más bien elemental. Quizás por este carácter elemental, los filósofos le han prestado poca o ninguna atención. A pesar de este desinterés, las *cortaduras de Dedekind* son, probablemente, el ejemplo más conocido de una completación.

La completación, utilizando las cortaduras de Dedekind, es particularmente sugerente, porque en ella se entrelazan aspectos geométricos y algebraicos de una manera fascinante.

Tomemos como punto de partida, los números racionales \mathbf{Q} . De nuevo, la línea euclidiana, proporciona un modelo geométrico de \mathbf{Q} : basándose

en el modelo geométrico de los números enteros, mencionados anteriormente, es intuitivamente claro, cómo pueden representarse todos los números racionales, como puntos de la línea euclidiana:



Naturalmente, no acertamos que todos los puntos de la línea euclidiana, corresponden a números racionales. Aunque los números racionales, son suficientes para realizar cálculos prácticos, no lo son, para aplicaciones de tipo más teórico. Por ejemplo, no son demasiado útiles para resolver ecuaciones polinomiales. En muchas ocasiones, se necesitan “más números”.

En otras palabras, el dominio de los números racionales, no ofrece un escenario lo suficientemente amplio como para resolver ecuaciones polinomiales. El ejemplo más clásico que muestra este defecto es la ecuación $x^2 = 2$. Para vencer esta deficiencia, debemos construir nuevos números “ideales”, que ofrezcan soluciones para ecuaciones, que no tienen soluciones en \mathbb{Q} .

Esto puede remediarse, gracias a la completación idealizadora de los números racionales \mathbb{Q} , a través de las cortaduras de Dedekind, que nos llevan a los números reales \mathbb{R} .

Por tal motivo, para construir los números que faltan, Dedekind tomó el conjunto de las “cortaduras” de los racionales. Una cortadura se basa en una idea muy intuitiva.

Una cortadura de Dedekind es una partición (I, S) de \mathbb{Q} , que cumple las siguientes condiciones:

$$I \cup S = \mathbb{Q}, I \cap S = \emptyset \ \& \ (x) (y) (x \in I \ \& \ y \in S \Rightarrow x < y)$$

Llamaremos a I *cortadura inferior* y a S *cortadura superior*. Obviamente, I y S se determinan la una a la otra:



Si denominamos el conjunto de las cortaduras inferiores de \mathbb{Q} , como $I(\mathbb{Q})$ se define una inclusión canónica de \mathbb{Q} , en $I(\mathbb{Q})$

$$\mathbf{Q} \xrightarrow{e} I(\mathbf{Q}), e(r) := \{x; x < r\}.$$

Trasladando el número racional r , a la cortadura inferior $e(r) := \{x; x < r\}$. De esta forma, los números racionales r , pueden identificarse con las cortaduras inferiores, que a su vez, definen cortaduras superiores $\{a; a \geq r\}$, con elementos mínimos r .

El punto crucial es que hay muchas *otras cortaduras inferiores*, cuyas cortaduras superiores, NO tienen elementos mínimos. Por ejemplo, $\{x; x^2 < 2\}$, la cual corresponde a la raíz cuadrada de 2.

Con base en esto, se muestra que la adición y multiplicación de los números racionales \mathbf{Q} pueden extenderse a $I(\mathbf{Q})$, esto es, $I(\mathbf{Q})$, se identifica con los números reales \mathbf{R} . A través, de la inclusión de \mathbf{Q} en $I(\mathbf{Q})$, el dominio \mathbf{Q} , de los números racionales, ha sido completado por ciertos "elementos ideales", de tal manera que $I(\mathbf{Q})$, es más cómodo desde el punto de vista de la aritmética, esto se muestra, por el hecho que ecuaciones como $x^2 = 2$ se pueden resolver en \mathbf{R} , pero no en \mathbf{Q} .

No obstante, tampoco \mathbf{R} es del todo satisfactorio, porque no todos los polinomios tienen solución o no tienen todas las soluciones que deberían tener, por ejemplo, la ecuación $x^2 + 2 = 0$, no tiene solución en \mathbf{R} , mientras que $x^3 - x^2 + 2x - 2 = 0$, sólo tiene una solución (y debería tener tres). Para conseguir una teoría de las ecuaciones polinomiales, completamente satisfactoria en la cual el teorema fundamental del álgebra sea válido, hay que sustituir el cuerpo incompleto \mathbf{R} y reemplazarlo por el cuerpo algebraicamente completo \mathbf{C} de los números complejos.

Podemos considerar los números complejos como parejas ordenadas (a, b) , de números reales a, b , tal que, un número real a se identifica con la pareja $(a, 0)$. De nuevo, este proceso, puede concebirse como un proceso de completación idealizante, sin embargo, de un tipo diferente al de la completación de Grothendieck o al de la completación basada en las cortaduras de Dedekind.

Un modelo geométrico bien conocido de los números complejos $\mathbf{C} := \{a + ib; a, b \in \mathbf{R}\}$, es el plano cartesiano, dotado con un sistema de coordenadas apropiadas, como el propuesto por Gauss: un punto (a, b) ($a, b \in \mathbf{R}$) del plano, corresponde de una manera 1-1, al número complejo $a + ib \in \mathbf{C}$.

Los cuerpos \mathbf{R} y \mathbf{C} , no son las únicas completaciones de \mathbf{Q} , quizás no son ni siquiera, las más interesantes matemáticamente. Hay completaciones intermedias, por decirlo así, que desempeñan un papel central en la teoría algebraica de los números y campos relacionados. Históricamente, el matemático alemán Kummer, parece haber sido el primero en considerar explícitamente, las completaciones de los números enteros ordinarios, mediante la introducción de ciertos números ideales.

En concreto, Kummer definió números primos ideales para cuerpos algebraicos $\mathbf{Q}[x]$, siendo x , una raíz primitiva de una ecuación ciclotómica. Por ejemplo, si tomamos la ecuación ciclotómica $x^3 = 1$, con la raíz primitiva $x = (-1 + \sqrt{3}i)/2$, se obtiene el cuerpo algebraico $\mathbf{Q}[x]$, $\mathbf{Q} \subseteq \mathbf{Q}[x] \subseteq \mathbf{C}$, cuyos “números enteros ideales” tienen la forma $(a + b(-1 + \sqrt{3}i)/2)$, $a, b \in \mathbf{Z}$.

Estos números enteros ideales, no deberían ser considerados como generalizaciones, más o menos bien motivadas, de los números enteros ordinarios, sino como elementos ideales que contribuyen a revelar la esencia de estos últimos:

[T]he new elements are not simply adjoined to the old ones as elements of a different kind and origin, but the new area systematically unfolding of the old. And this requires... primary logical kinship between the two. Then the new elements will bring nothing to the old, other than what was implicit in their original meaning. If this is so, we may expect that the new elements, instead of fundamentally changing the meaning and replacing it by another, will first bring it to its full development and clarification. And when we survey the history of the ideal elements in mathematics, this expectation is never disappointed.

[...] [W]e have to use the new elements as an intellectual medium by which to apprehend the true meaning of the old, by which to know it with universality and depth never before achieved. (Cassirer 392)

Los matemáticos del siglo XIX, entendieron estos elementos ideales de una manera muy similar a Cassirer. Por ejemplo, en 1859, el matemático irlandés H. J. S. Smith, defendió en su *Report on the theory of numbers*, en la Royal Society:

Los números complejos de Gauss, Jacobi y Kummer se imponen a nuestra consideración, no porque sus propiedades sean generalizaciones de las propiedades de los números ordinarios, sino porque los números enteros sólo pueden explicarse en una referencia a ellos. (Smith 1860)

Dicho de otra manera, la interpretación de Cassirer del papel de las idealizaciones, sigue la interpretación que ya había sido ofrecida por diversos matemáticos. Esta interpretación, implica que los objetos de las teorías matemáticas, no se conciben como algo dado o ya “fijado *a priori*”. Al contrario, estos objetos, pueden ser sujetos a construcciones idealizantes. Estas construcciones, imposibilitan una concepción de la matemática como una ciencia de lo intuitivamente dado, es decir, como una ciencia de un dominio platónico invariable.

Completación de álgebras booleanas

Uno de los propósitos fundamentales del presente artículo, es el de convencerles de que el concepto de “completación idealizante”, juega un papel central en la matemática “real”, es decir, las matemáticas de los matemáticos y que no es una mera invención de los filósofos. Empecemos con un teorema, este es considerado por los matemáticos como uno de los teoremas más importantes del siglo pasado³, el famoso “teorema de representación de Stone”, Davey & Priestley (1990), Grosholz (1985), Johnstone (1982).

Este teorema, estableció el primer vínculo no-trivial, entre la topología y la lógica. Fue demostrado por el matemático americano Marshall Stone, a mediados de los años 30. El teorema de Stone, demuestra que cada álgebra booleana B , tiene una *representación topológica*, por medio del álgebra booleana *completo* B^* . Podemos encontrar ejemplos familiares de álgebras booleanas en los conjuntos potencia PX de X o bien, en las álgebras de Lindenbaum de la lógica proposicional⁴.

De hecho, el teorema de representación de Stone, es una generalización de la completación de Dedekind, de los números racionales \mathbf{Q} , a los reales \mathbf{R} . Más precisamente, el álgebra B , corresponde a los números racionales \mathbf{Q} y su completación B^* , corresponde a los números reales \mathbf{R} .

³ Véase el libro clásico *Stone spaces*, de Johnstone (1982).

⁴ Una demostración elemental, pero completa, del teorema de Stone, se encuentra en el libro de Davey & Priestley (1990).

De una manera más detallada, B^* se define como el álgebra completa de los *subconjuntos abiertos de tipo regular* de un espacio topológico, llamado espacio Stone (B). Los puntos del espacio topológico Stone (B) son los ideales primos del álgebra booleana B con la que hemos empezado.

La completación de Stone de álgebras booleanas, por medio de ideales en el sentido de la teoría de los retículos, abrió el camino para establecer muchos vínculos entre el álgebra y la topología (Grosholz). Además, la completación de Stone se convirtió en el prototipo de completaciones en muchas otras áreas de la matemática (Johnstone).

Una aplicación interesante del teorema de Stone, está en la base de la aplicación de la matemática, a las ciencias empíricas como la física. Esto tiene que ver, con los puntos como construcciones ideales de la geometría física. Este método fue esbozado por Whitehead (1929), en su libro *Process and reality*, pero un tratamiento matemáticamente más preciso, requiere a su vez del teorema de Stone.

Esencialmente, el problema subyacente puede describirse de la siguiente manera: aunque en física, geometría y otras ciencias se habla frecuentemente de los puntos como elementos atómicos básicos, en la "realidad", nunca encontramos puntos. Consideremos la realidad física. Habitualmente, el físico no siente escrúpulos al hablar de puntos espaciales o acontecimientos instantáneos, interpretados como puntos temporales, entre otros, no obstante, también admitirá, sin problemas, que nunca experimenta de hecho, estas entidades llamadas puntos. Por ejemplo, los valores medidos de una magnitud física, nunca se suponen exactos, simplemente porque esto equivaldría a suponer una precisión absoluta que no se da en la ciencia real. Por ende, se acostumbra a decir, que los valores medidos se localizan en cierto intervalo, pero habitualmente, dicho intervalo se entiende como un conjunto de puntos.

Así, aunque se admita que nuestros métodos de medición nunca alcanzan la exactitud de los puntos, estamos acostumbrados a considerar a éstos como las entidades básicas del espacio, el tiempo y el espacio-tiempo, por más que desde un punto de vista estrictamente empírico, parezcan ser entidades bastante dudosas.

Whitehead, fue uno de los primeros en intentar sustituir los puntos como entidades fundamentales por otras entidades empíricamente más accesibles. En el caso del espacio y el espacio-tiempo, estas entidades, pueden denominarse *regiones*. Intuitivamente, una región espacial, se

puede entender, como una parte del espacio, con límites más o menos precisos.

El programa de Whitehead consistía en considerar las regiones, en lugar de los puntos como entidades básicas fundamentales y construir los puntos y sus relaciones geométricas a partir de las regiones y sus relaciones. Whitehead sólo presentó un esbozo informal de cómo podía funcionar esto, pero se puede demostrar que su planteamiento es susceptible de una formalización rigurosa.

Lo más interesante es que dicha construcción de los puntos, a partir de las regiones, puede entenderse como una generalización del método de las “cortaduras” de Dedekind; es decir, que la introducción de los puntos como elementos ideales o límite en el ámbito de las regiones es un proceso idealizante de compleción, análogo a la construcción de los números reales a partir de los racionales. Con mayor precisión, esta compleción, se entiende como una generalización del método de construcción que utilizó Stone en su demostración del teorema de representación, que lleva su mismo nombre.

Así mismo, la afirmación que los métodos de compleción al estilo de Dedekind, forman parte de las síntesis fundamentales más importantes sobre las que se basan las matemáticas y las ciencias empíricas, queda más justificada, cuanto más nos introducimos en los vericuetos de la matemática y la ciencia moderna.

Resumiendo, podemos decir que el método de los elementos ideales es un procedimiento poderoso, presente en muchas disciplinas matemáticas de la geometría física y otras ciencias. Los elementos ideales que introduce, no son exactamente otro tipo de cosas, que añadimos a las “reales”, sino que, más bien, expresan cierta forma de tratar las cosas “reales”. Por ejemplo, una cortadura de Dedekind, es el límite de una acción ideal, la de repartir el conjunto de los números racionales en dos mitades; esta acción ideal, no se supone que sea posible en la realidad, debido a que es una especie de supertarea, que equivaldría, a la determinación de los números infinitos con precisión absoluta.

Las acciones empíricamente posibles, sólo apuntan en la dirección de la cortadura ideal, como dijo Cassirer. Esto es exactamente, lo que Cassirer pretendía subrayar, cuando afirmaba que, las leyes exactas,

matemáticamente formuladas, no se cumplen en el terreno de nuestras experiencias sensibles, sino tan solo, para objetos límites ideales:

Si se pretende que tenga éxito la fundamentación de la teoría exacta de la cinemática, el cuerpo rígido de la geometría pura tiene que sustituir a los cuerpos perceptibles y su variabilidad infinita.

[...] Razonablemente, las leyes sólo se pueden formular para objetos límites ideales que sustituyen conceptualmente a los objetos de la percepción empírica. "Estar en movimiento" es un predicado que nunca se puede aplicar a las "cosas" del mundo sensible que nos rodea. Sólo se cumple para ese otro tipo de objetos que el matemático pone en su lugar en una construcción libre. El movimiento no es una cuestión de sensación sino de pensamiento, no de percepción sino de concepción. (Cassirer 118)

El papel de las idealizaciones es esencialmente, el mismo en todos los terrenos. Son necesarias para asegurar la existencia de fenómenos estables, llámense leyes en la física o teoremas en la matemática. El papel de los conceptos límite, es el mismo para la constitución de campos de números, que para dominios geométricos o empíricos. Así mismo, uno de los principales objetivos de la ciencia es la construcción de "situaciones estables" que muestran cierto tipo de comportamiento invariante, "sujeto a leyes" en un océano de contingencias.

En los apartados anteriores, he caracterizado los medios para construir tales situaciones estables, con varios métodos, al introducir los elementos límites idóneos. Estos métodos, se inspiran en la matemática, pero métodos similares, se emplean también en las ciencias empíricas: un caso destacable, sería la construcción de los puntos en la geometría física.

La completación de Yoneda

El segundo ejemplo de una completación importante en la matemática contemporánea es el "lema de Yoneda". Este lema, trata con completaciones de categorías. De manera muy breve, la noción de categoría, se describe como un universo local para el discurso matemático. Por ejemplo, la categoría de grupo GRP o categoría de

espacios topológicos TOP, definen el marco de objetos y relaciones, que tratan con los grupos y los espacios topológicos respectivamente.

Las categorías tienen ciertos rasgos estructurales que las hacen más o menos apropiadas para llevar a cabo ciertas “construcciones universales”. A veces, la categoría en la cual estamos trabajando, no nos permite llevar a cabo algunas de estas construcciones universales. Esta dificultad es similar al problema que se ha discutido al inicio, es decir, el problema de solucionar ciertas ecuaciones en ciertos sistemas de números.

Por tanto, en muchas ocasiones es útil reemplazar la categoría original, por una completación más apropiada. El lema de Yoneda, nos ofrece un método versátil de completación. Éste, ha resultado de gran utilidad en muchas áreas distintas de la matemática. Grosso modo, una categoría A “localmente pequeña”, se incluye por el functor de Yoneda Y , en una categoría functor

$$Y : A \longrightarrow \text{SET}^{A^{\text{opp}}}$$

Aquí, SET es la categoría de los conjuntos y A^{opp} es la categoría “opuesta” de A . Esta última categoría, puede considerarse como una especie de imagen espejo de A . La relación entre A y A^{opp} es análoga a la relación entre “+1” y “-1”. Cada morfismo $a \longrightarrow b$, de la categoría A , corresponde biunívocamente, a un morfismo $b \xrightarrow{f^{\text{opp}}} a$. En particular $A^{\text{oppopp}} = A$, analógicamente como $(-1)^2 = +1$. A pesar de su carácter elemental, la inclusión de Yoneda es una construcción con una amplia gama de aplicaciones. En un libro de texto contemporáneo de la teoría de las categorías, encontramos la siguiente observación:

El lema de Yoneda es quizás el resultado más utilizado en la teoría de categorías...

[...] La categoría $\text{SET}^{A^{\text{opp}}}$ es como una extensión de A por “elementos ideales” que permite cálculos que no se puede hacer en A . Esto es algo así como pasar a los números complejos para resolver las ecuaciones de los reales, o la adición de un tipo más elevado de una teoría lógica elemental. (Awodey 194)

La tesis de Awodey, que se encuentra en una u otra forma, en muchos libros de texto, me parece particularmente relevante para confirmar

nuestro enfoque de la idealización como completación, porque en este libro, su autor no tiene ningún interés en asuntos filosóficos. Como dije antes, es un libro de texto sobre teoría de *las categorías* para los matemáticos.

Para dar un ejemplo concreto de una aplicación del lema de Yoneda quisiera mencionar “el análisis infinitesimal suave (Smooth Infinitesimal Analysis)”. Básicamente este “análisis suave” es algo así como la resurrección bastante sorprendente, de los infinitesimales de la matemática pre-moderna.

En la mayoría de las versiones pre-modernas del “cálculo infinitesimal”, infinitesimales dx , dy , esto es, magnitudes infinitamente pequeñas, juegan un rol esencial.

En las últimas décadas del siglo XIX, los infinitesimales, fueron expulsados del discurso oficial de la matemática, gracias a los trabajos de matemáticos como: Cantor, Dedekind y Weierstraß.

Filósofos como Carnap y Russell consideraron los infinitesimales como “seudo-conceptos” metafísicos que deberían ser eliminados del discurso matemático. Sorprendentemente, en los años 70 del siglo pasado, se demostró, que números reales \mathbf{R} , podían completarse de tal manera que el sistema resultante, contenía infinitesimales. En este proceso de completación, el lema de Yoneda juega un papel crucial. Nos permite completar los números reales de toda la vida, con infinitesimales no-arquimedianos. Así, la estructura resultante, sirve como base de un análisis infinitesimal suave (Smooth Infinitesimal Analysis).

Consideraciones finales

Como he dicho anteriormente, mi propósito era convencerles de que las completaciones juegan un papel importante en la evolución conceptual de las matemáticas y que no son una quimera inventada por filósofos. Además, estas completaciones no están reducidas a partes demasiado elementales de la matemática.

Así mismo, cualquier filosofía de las matemáticas, con algún interés en las matemáticas reales y que no se contente con modelos filosóficos de la matemática más o menos anémicos, debería reconocer la importancia de

esta herramienta conceptual. Si las completaciones son importantes para la matemática, sería natural preguntarse, si juegan un papel similar en las ciencias empíricas. Después de todo, las idealizaciones en las ciencias empíricas, están basadas en muchas ocasiones en consideraciones matemáticas.

Cassirer contestó esta cuestión con la “tesis de identidad” (TI), con un claro y rotundo Sí:

(TI) la idealización en la matemática y las ciencias empíricas matematizadas, puede ser caracterizada mediante la introducción de “elementos ideales”. En ambas áreas, la introducción de estos elementos ideales, juega esencialmente, el mismo papel.

Cassirer que entendió la “ciencia” en el sentido amplio del concepto alemán “Wissenschaft” enfatizó que, *tanto* en matemáticas, *como* en física, se utiliza el método de los elementos ideales. Esta afirmación, contrasta fuertemente con las opiniones actuales, según las cuales, la matemática está por decirlo así, *dentro* de la esfera de los objetos ideales, mientras que la física, queda fuera de ese mundo ideal.

En contraste con la mayoría de los estudios actuales, para Cassirer no existía una separación tajante entre las matemáticas y las ciencias empíricas. Lo que afirma Cassirer es que la idealización, esto es, la introducción de elementos ideales, es tan común a la matemática, como a la física.

Así, la matemática, no debería caracterizarse únicamente como un ámbito de objetos ideales y tampoco la física debe limitarse a la esfera empírica. En principio, la matemática como ciencia, no es tan diferente de la física.

Cassirer sostuvo la tesis (ST), durante toda su carrera filosófica. Quizá, la versión más elaborada, la encontramos en su obra maestra *La filosofía de las formas simbólicas* (1955). Aquí, Cassirer propuso (ST), como una formulación alternativa de su “principio fundamental de la unidad de la ciencia”:

La lógica, las matemáticas y la física no deben ser concebidas como diferentes dominios ontológicos que existían por

separado. Más bien, debemos comparar la matemática, la lógica y las formas físicas de conocimiento en su evolución histórica en curso. Esta comparación nos llevaría a la conclusión de que ninguna de estas formas puede constituir un ser objetivo y el ámbito de la validez objetiva teórica para sí misma, sino que todos ellos se construyen juntos en sus interconexiones mutuas. (PSF III 384 ff.)

No todo el mundo compraría esta afirmación. Pero al menos, podríamos considerarla como un argumento a favor de una tesis más débil, pero más plausible, es decir, la tesis de la similitud:

(TS) una teoría filosófica comprensiva de la idealización, tiene que tomar en cuenta las idealizaciones, tanto en matemáticas, como en las ciencias. En ambas esferas, la introducción de elementos ideales, juega un rol similar.

Finalmente, sostengo que esta tesis de *la similitud*, entre idealización en ciencias y en matemáticas, puede ser la máxima de un programa de investigación prometedor.

REFERENCIAS BIBLIOGRÁFICAS

Avidad, Jeremy. "Methodology and metaphysics in the development of Dedekind's theory of ideals". *The architecture of mathematics*. Eds. Jose Manuel Ferreirós and Jeremy Gray. Oxford: Oxford University Press, 2006. Print.

Awodey, Steve. *Category theory*. Oxford: Oxford University Press, 2006. Print.

Beiser, Frederick C. "Kant's intellectual development: 1746-1781". *The Cambridge companion to Kant*. Ed. Paul Guyer. Cambridge: Cambridge University Press, 1992. Print.

Cassirer, Ernst. "Kant und die moderne mathematik". *Kantstudien*. 1907: 1-49. Print.

---. "Das Erkenntnisproblem in der Philosophie und Wissenschaft der neueren Zeit". *Die nachkantischen Systeme*. 1920. Print.

---. *The philosophy of symbolic forms*. New Haven: Yale University Press, 1955. Print.

---. *Substance and function & Einstein's theory of relativity*. New York: Dover, 1953. Print.

Corry, Leo. *Modern algebra and the rise of mathematical structures*. Vol. 17 of *Science Networks*. Basel: Birkhäuser Verlag, 1996. Print.

Corfield, David. *Towards a philosophy of real mathematics*. Cambridge: Cambridge University Press, 2003. Print.

Davey, Bryan. and Hilary A. Priestley. *Introduction to lattices and order*. Cambridge: Cambridge University Press, 1990. Print.

Grosholz, Emily. "Two episodes in the unification of logic and topology". *British Journal of Philosophy of Science*. 1985: 147-157. Print.

Johnstone, Peter. *Stone space*. Cambridge: Cambridge University Press, 1982. Print.

Mac Lane, Saunders. *Mathematics. Form and Function*. New York and Berlin: Springer, 1986. Print.

Mormann, Thomas. "Idealization in Cassirer's Philosophy of Mathematics". *Philosophia Mathematica*. 2008: 151-181. Print.

Norton, John D. "Approximation and idealization: Why the difference Matters". *Philosophy of Science*. Forthcoming.

Nowakowa, Izabella. and Leszek Nowak. "The richness of idealization". *Poznan studies in the philosophy of the sciences and the Humanities*. 2000. Print.

Sieg, Wilfred and Dirk. Schlimm. "Dedekind's analysis of number: Systems and axioms". *Synthese*. 2005: 121-170. Print.

Weisberg, Michael. "Three kinds of idealization". *The Journal of Philosophy*. 2007: 639-659. Print.

Whitehead, Alfred North. *Process and reality: An essay in cosmology*. New York: MacMillan, 1929. Print.

UNA DEFENSA DE LA METAINDUCCIÓN PESIMISTA

A DEFENSE OF THE PESSIMISTIC METAINDUCTION

DAIAN TATIANA FLÓREZ QUINTERO

Universidad de Caldas, Colombia, daian.florez@ucaldas.edu.co

Universidad Nacional de Colombia - Sede Manizales

RECIBIDO EL 22 DE MARZO DE 2012 Y APROBADO EL 24 DE MAYO DE 2012

RESUMEN ABSTRACT

En el presente artículo, me propongo ofrecer una defensa del argumento 'metainductivo' propuesto por Laudan (1981) en contra del realismo epistemológico convergente (en adelante REC). Para tal fin, analizaré dos de las líneas de ataque que se han dirigido en contra del argumento metainductivo: la primera de ellas, ofrecida por Antonio Diéguez (1998), quien acusa a Laudan de atribuir ilegítimamente un conjunto de premisas ocultas a los argumentos realistas, las cuales no son mantenidas, ni implícita, ni explícitamente por los defensores del realismo. La segunda línea de ataque es dirigida por Stathis Psillos (1999), quien intenta reducir el número de contraejemplos ofrecidos por Laudan para mostrar que la base inductiva de su argumento no es tan amplia, ni lo suficientemente representativa como para garantizar la inducción pesimista. Adicionalmente, Psillos pretende mostrar que hay serias dudas de que todas las teorías de Laudan pertenezcan a un estadio de desarrollo lo suficientemente maduro o que hubiesen sido exitosas.

Para responder al primer ataque, haré un contraste entre los supuestos atribuidos por Laudan al realista y aquellos que Diéguez considera que Laudan atribuye de manera ilegítima al REC. En cuanto a los ataques de Psillos, mostraré que buena parte de su línea argumentativa ha sido anticipada por Laudan, quien logra responder de manera categórica a dicho ataque. Por estas razones, no parece persuasivo tratar de reducir la lista y señalar que buena parte de las teorías mencionadas en el argumento no pertenecen a estadios de desarrollo maduro de la ciencia, porque -tal y como lo indica Laudan- una estrategia como ésta estaría condenada al fracaso, dado que el realista tendría que admitir que no dispone de una explicación del éxito general de la ciencia, independientemente del estadio en que se encuentren las teorías.

In this paper, I shall offer a defense of Laudan's 'metainductive' argument (1981) against Convergent Epistemological Realism (hereafter REC). To this end, I shall analyze two lines of attack that have been directed against the metainductive argument: the first one, offered by Antonio Dieguez (1998), charges Laudan with adding illegitimately a set of hidden premises to the realist arguments, which are not maintained, neither implicitly nor explicitly, by advocates of realism. The second line of attack, directed by Stathis Psillos (1999), tries to reduce the number of counterexamples offered by Laudan to show that the inductive basis of his argument is not broad enough, nor sufficiently representative to guarantee pessimistic induction. Additionally, Psillos tries to show that there are serious doubts about Laudan's theories belonging to a developmental stage mature enough or that they have been successful.

To answer the first attack, I shall make a contrast between the assumptions attributed by Laudan to the realist, and those that Dieguez thinks are illegitimately attributed by Laudan to REC. Regarding Psillos' attacks, I shall show that much of his argumentative line has been anticipated by Laudan, who answers successfully such attack. For this reason, an attempt at reducing the list, indicating that most of the theories mentioned in the argument do not belong to mature stages of science's development does not seem persuasive, because -as indicated by Laudan- such a strategy would fail, given that the realist would have to admit that he does not have an explanation of the success of science, regardless of the stage in which theories are.

PALABRAS CLAVE KEY WORDS

Argumento metainductivo, Diéguez, éxito, Psillos, realismo epistemológico convergente, referencia.

Metainductive argument, success, Dieguez, Psillos, convergent epistemological realism, reference.

* Este artículo, hace parte del proyecto de investigación: "¿Es posible una defensa sistemática del antirrealismo?", financiado por la Vicerrectoría de Investigaciones y Postgrados, de la Universidad de Caldas.

Supuestos del realismo convergente

En el presente acápite indicaré las tesis, los argumentos y los supuestos que Laudan atribuye al REC, deteniéndome en el examen de los supuestos que Diéguez considera que son atribuidos ilegítimamente por Laudan a los defensores del realismo. Con base en esto, mostraré que sólo dos de los tres supuestos identificados por Diéguez, son atribuidos explícitamente por Laudan al REC.

Tesis, argumentos y supuestos del REC

Según Laudan, el realismo que él critica, incluye variantes de las siguientes afirmaciones:

(R1) Las teorías científicas (al menos en las ciencias 'maduras'), son aproximadamente verdaderas y las teorías más recientes están más cerca de la verdad, que sus predecesoras en el mismo dominio.

(R2) Los términos observacionales y teóricos de las teorías de una ciencia madura, refieren genuinamente (aproximadamente, hay sustancias en el mundo, que corresponden a las ontologías presumidas por nuestras mejores teorías).

(R3) Teorías sucesivas en una ciencia madura serían tales que preservan las relaciones teóricas y los referentes aparentes de las teorías predecesoras; es decir, las teorías predecesoras serían casos límite de las teorías más recientes.

(R4) Las nuevas teorías aceptadas explican y deberían explicar por qué sus predecesoras fueron exitosas.

(R5) Las tesis (R1) a (R4) implican que las teorías científicas ("maduras") deberían ser exitosas; es más, estas tesis, constituyen la mejor, si no la única explicación para el éxito de la ciencia (Laudan 224).

Estas tesis son defendidas por los realistas, según Laudan, a partir de los siguientes dos argumentos abductivos:

1er Argumento abductivo:

(P1) Si las teorías científicas son aproximadamente verdaderas, entonces son empíricamente exitosas.

(P2) Si los términos centrales de las teorías científicas refieren genuinamente, entonces tales teorías son empíricamente exitosas.

(P3) Las teorías científicas son empíricamente exitosas.

Por lo tanto, (probablemente) las teorías son aproximadamente verdaderas y sus términos refieren genuinamente (Laudan 225).

2do Argumento abductivo:

(P1) Si las teorías predecesoras en una ciencia 'madura' son aproximadamente verdaderas y sus términos refieren genuinamente, entonces las teorías más exitosas de la misma ciencia deberían preservar a las teorías predecesoras como casos límite.

(P2) los científicos se proponen preservar a las teorías predecesoras como casos límite y generalmente tienen éxito al hacerlo.

Por lo tanto, (probablemente) las teorías predecesoras de una ciencia 'madura' son aproximadamente verdaderas y genuinamente referenciales (Ibíd. 25).

A la base de estos argumentos, Laudan identifica las siguientes cuatro premisas ocultas las cuales, a su juicio, son significativamente problemáticas por lo que no se pueden reconciliar con el registro histórico de la ciencia.

(S1) Las teorías en las ciencias avanzadas o maduras son exitosas.

(S2) Una teoría cuyos términos centrales refieren genuinamente debe ser una teoría exitosa.

(S3) Si una teoría es exitosa, entonces sus términos refieren genuinamente.

(S4) Todos los términos centrales de las teorías científicas en las ciencias maduras refieren genuinamente (Laudan 225).

De acuerdo con Laudan (S2) y (S4), explican a (S1) y a su vez (S1) y (S3) son garantía de (S4). De lo que se sigue que la referencia explica el éxito y el éxito es una garantía para la presuposición de la referencia. Lo que Laudan propone mostrar a partir de su análisis, es que no hay una conexión legítima entre la referencia y el éxito de suerte que no es evidente que la referencia desempeñe un rol explicativo causal en el realismo científico.

El rechazo de la conexión entre referencia y éxito

(S1) es inadecuada, porque el uso que se hace del término 'éxito' es ambiguo. De modo que para poder llevar a cabo una interpretación caritativa de las tesis realistas, Laudan parte de la suposición según la cual después de todo el realista está principalmente interesado en saber por qué la ciencia ha trabajado tan bien. Por lo que se podría asumir que el uso del término 'éxito' es de carácter pragmático, así que es correcto afirmar que una teoría es exitosa *sys*: i) ha funcionado en una variedad de contextos explicativos; ii) ha permitido hacer predicciones bien confirmadas; iii) ha permitido la ampliación de su campo explicativo.

Si asumimos que las tres condiciones anteriores indican en qué consiste el éxito de la ciencia, entonces (S1) no es una premisa problemática. Sin embargo, si las condiciones del éxito fueran: i) que las teorías científicas refieran genuinamente, i.e. que las entidades postuladas existan y ii) que la mayoría de sus enunciados sean verdaderos; dichas condiciones se vuelven problemáticas al ser sometidas al test de la experiencia. Esto es lo que Laudan muestra a partir de su análisis de (S2).

(S2) es inaceptable, porque a pesar de que el sentido realista de la referencia, es liberal (dado que, de acuerdo con los realistas, los términos de una teoría, deberían referir aún si muchas de sus afirmaciones fuesen falsas), según Laudan, hay en principio tres contraejemplos que muestran la falsedad de la afirmación según la cual: *si los términos centrales de una teoría refieren genuinamente, entonces la teoría es exitosa*. Por ejemplo, la teoría atómica durante el siglo XVIII, la teoría de Prout y la teoría del desplazamiento continental de Wegener¹.

¹ Se puede afirmar legítimamente que dicha teoría fracasa porque, tal y como lo indica Doppelt, no permitía hacer nuevas predicciones, ni ofrecía explicaciones de nuevos fenómenos a pesar de que podía explicar una amplia variedad de fenómenos geológicos, paleoclimáticos y paleomagnéticos. Por ejemplo, podía explicar la forma de los continentes y la existencia de fauna y flora fósil similar a ambos lados del Atlántico; además, proporcionaba evidencia de que los polos se habían movido, entre otras (Doppelt 105).

Estos contraejemplos rebaten felizmente la verdad de dicha implicación, dado que las tres teorías refieren genuinamente, i.e, estipulan entidades que existen, tales como: 'átomo' y 'manto terrestre' y sin embargo no pueden ser estimadas como teorías genuinamente exitosas debido a que fueron ineficaces en la solución de problemas, lo que se puede probar a partir del hecho de que los químicos del siglo XVIII prefirieron la teoría de la afinidad electiva en lugar de la teoría química. La teoría de Wegener, por su parte, requirió de treinta años y de modificaciones muy importantes en su base, para que rindiera frutos en las explicaciones geológicas (Laudan 227).

A la luz de una interpretación caritativa, Laudan se propone formular una versión deflacionaria de (S2) en los siguientes términos: (S2') *una teoría cuyos términos refieran sería (aunque no necesariamente siempre) exitosa*. La principal dificultad que enfrenta (S2') consiste en: si encontramos un conjunto significativamente amplio de teorías que refieren genuinamente, pero cuya falta de éxito sea evidente, deberíamos rechazarla.

En contra de la versión debilitada de (S2), Laudan afirma que, consideremos cuántas versiones inadecuadas de la teoría atómica se ofrecieron durante dos mil años de especulación sobre el tema antes de que surgiera una genuinamente exitosa y cuántas versiones sobre la teoría ondulatoria de la luz se formularon sin ningún éxito, antes de emerger una genuinamente exitosa en 1820. Tales casos prueban que (S2') es aún más difícil de reconciliar con los registros históricos. Ahora bien, si estos ejemplos no bastan podemos considerar las teorías cinéticas del calor durante los siglos XVII y XVIII y las teorías embriológicas antes de finales del siglo XIX (Ibíd. 227).

Por otra parte, a pesar de que (S3) no se enfrenta a las objeciones anteriores tiene sus propias dificultades. En primer lugar, parece difícil de coordinar con el hecho de que los términos centrales de teorías relativamente exitosas no refieran genuinamente, como las teorías del éter y la teoría del flogisto. No obstante, si consideráramos una versión deflacionaria de (S3) que en lugar de afirmar que si la teoría es exitosa, entonces sus términos refieren genuinamente, estableciera: (S3') *la mayoría de las teorías exitosas refieren genuinamente*, nos encontraríamos con una versión que tampoco está exenta de dificultades. Su principal defecto estriba en que esta premisa podría ser o no verdadera, por lo que no puede conferírsele el rol explicativo a la referencia, tal y como el realista lo desea.

A la luz del análisis de Laudan sobre el punto de vista anterior, se sigue que el realista pretende establecer una conexión entre éxito y referencia sobre la presuposición de que las teorías que son exitosas son aproximadamente verdaderas. El razonamiento discurriría de la siguiente manera:

*Si una teoría es exitosa, entonces es aproximadamente verdadera
Si es aproximadamente verdadera, entonces refiere genuinamente*

Por tanto, una teoría exitosa refiere genuinamente.

El reparo de Laudan a un argumento de esta naturaleza, consiste en sostener que si bien el vínculo entre referencia y éxito se establece a través de la aproximación a la verdad, dicho medio es extremadamente tenue. En contra de la conclusión del argumento anterior, Laudan ofrece la siguiente evidencia histórica:

Argumento histórico a partir de las teorías del éter: de acuerdo con Laudan, la presuposición de la existencia del éter permitió explicar fenómenos muy disímiles entre sí, con un notable éxito, a pesar de que los experimentos llevados a cabo por Michelson-Morley revelaron que el éter no existía. Por ejemplo, la presuposición del éter eléctrico permitió explicar la atracción entre cuerpos con cargas opuestas. En el dominio de la química la presuposición del éter permitió explicar el rol del calor en las reacciones químicas, la conducción y la radiación entre otros problemas de termometría. En la teoría de la luz, el éter óptico funcionaba en las explicaciones sobre la reflexión, la refracción, la interferencia, la doble refracción, la difracción y la polarización (Laudan 228).

Podríamos considerar que una versión debilitada de (S3) es inmune a estos contraejemplos; sin embargo, esta premisa enfrenta la siguiente dificultad:

Si sólo algunos de los términos centrales de las teorías científicas refieren genuinamente, entonces no es posible mantener (R3)².

Del análisis anterior Laudan concluye que aún cuando el realista está dispuesto a conceder que algún conjunto no especificado de términos de una teoría exitosa podría no referir, entonces sus propuestas para

² Según Laudan, éste ha sido uno de los pilares del realismo por décadas.

restringir ‘la clase de las teorías candidatas’ a aquellas que preservan la referencia para los términos referenciales *prima facie* en teorías, carece de fundamento (Laudan 230). En lo que resta de este acápite, contrastaré la naturaleza de las premisas ocultas identificadas por Laudan y los ‘supuestos ilegítimos’ identificados por Diéguez.

Contraste entre los supuestos: Laudan vs Diéguez

Los ‘supuestos ilegítimos’ identificados por Diéguez son:

- 1) Una teoría cuyos términos no teóricos no refieren, no puede tener éxito jamás.
- 2) Una teoría cuyos términos teóricos refieren, tiene éxito siempre.
- 3) El éxito de una teoría justifica su verdad.

Una análisis detenido de (1) revela que Laudan no atribuye ningún supuesto al realista, con base en el cual estipule que si las teorías científicas no refieren, entonces no pueden ser exitosas. Esta es más bien una afirmación implicada por el rechazo de Laudan de los dos planteamientos siguientes: i) si las teorías refieren, entonces son exitosas y ii) si las teorías son exitosas, entonces refieren.

La atribución de Laudan al realista de estos dos supuestos está motivada en la declaración explícita de los realistas (Boyd (1973) Putnam (1975)), según la cual la referencia desempeña un rol explicativo causal del éxito de la ciencia. Al examinar la evidencia histórica disponible, Laudan encuentra que la ciencia ha exhibido importantes casos en los que teorías genuinamente referenciales carecen de éxito y de teorías que, sin ser referenciales, pueden ser estimadas como genuinamente exitosas. A la luz de semejantes contraejemplos, es perfectamente legítimo inferir (1), por medio del *Modus Tollens*.

Ahora bien, Diéguez sostiene que la atribución de (1) es ilegítima debido a que, lo que el realista está afirmando es que la referencia genuina, explica el éxito instrumental de las teorías; pero en modo alguno el realista sostiene que el éxito lo autorice a inferir que dicha teoría tiene que ser genuinamente referencial (Diéguez 108).

Semejante afirmación me parece sorprendente. Consideremos cualquier caso de explicación ordinaria en la ciencia; por ejemplo, aquella que estipula que el oxígeno es el principal agente causal en la combustión

de los cuerpos. A la luz de una explicación como ésta, se establece que el oxígeno explica por qué los cuerpos arden; pero si seguimos el razonamiento de Diéguez, no estaríamos autorizados a inferir, a partir del hecho de que un trozo de madera arda, que lo hace gracias a la intervención del oxígeno.

Avancemos ahora al análisis del segundo supuesto que Diéguez considera ilegítimo. De acuerdo con Diéguez el supuesto (2) estipula que si los términos teóricos de una teoría científica refieren, entonces es exitosa siempre. Este es un supuesto bastante llamativo, debido a que éste sí es atribuido por Laudan a los defensores del realismo. La dificultad que aquí subyace estriba en que el propio Laudan ha examinado una versión caritativa de dicho supuesto -precisamente a la que Diéguez se suscribe-, pero la encuentra tan insostenible como una versión fuerte.

Esto es así porque si el realista sostiene de manera débil que, *una teoría cuyos términos refieran sería (aunque no necesariamente siempre) exitosa*; tendría que enfrentarse al hecho de que ésta versión deflacionaria, no sólo resulta más difícil de reconciliar con el registro histórico de la ciencia, sino que adicionalmente, si establecemos semejante restricción, i.e, si mantenemos que no necesariamente la referencia explica el éxito de la ciencia, la referencia no podría servir como un genuino agente causal del éxito científico. La conclusión queda más clara si consideramos que en las explicaciones causales los agentes causales son -en términos lógicos- las *condiciones necesarias* de determinado fenómeno.

Por otro lado, el examen del tercer supuesto 'ilegítimo' según Diéguez, conduce a las razones ofrecidas por Laudan para rechazar la legítima conexión entre verdad y éxito, asunto del que pasaré a ocuparme enseguida.

Verdad aproximada y éxito

Tras concluir que no es legítimo el vínculo que el realista quiere establecer entre referencia y éxito, Laudan analiza si ocurriría lo mismo entre la aproximación a la verdad y el éxito. Este análisis procede mediante dos vías: *las vías del descenso y el ascenso*.

Con base en la primera vía, Laudan considera que si el realista quiere mantener que la aproximación a la verdad es el otro agente causal que explica el éxito de la ciencia, tendría que suscribirse a las siguientes tesis:

(T1) Si una teoría es aproximadamente verdadera, entonces sería explicativamente exitosa.

(T2) Si una teoría es explicativamente exitosa, entonces es probablemente, aproximadamente verdadera.

En contra de (T1), Laudan ofrece el siguiente argumento:

Supongamos, por ejemplo, que afirmamos en una línea popperiana, que T1 es aproximadamente verdadera si su contenido de verdad es mayor que su contenido de falsedad, es decir:

$$\text{Ctv (T1)} \gg \text{Ctf (T1)}.$$

Sin embargo, cuando la aproximación a la verdad es determinada, no se sigue lógicamente, que una clase arbitrariamente seleccionada de enunciados (a saber, las consecuencias observables), sea verdadera. Es más, es perfectamente concebible, que una teoría fuese aproximadamente verdadera en el sentido indicado, mientras todas sus consecuencias son falsas.

De este análisis, Laudan colige que los realistas no disponen de una teoría de la verosimilitud articulada, por lo que uno de los desafíos que plantea, consiste precisamente en exigir al realista una caracterización clara de lo que significa la locución 'aproximación a la verdad' (Laudan 231).

Por contraste, en la vía del ascenso, Laudan se propone examinar si es legítima o no, la conexión entre 'referencia genuina' y 'aproximación a la verdad', sosteniendo que: *ningún realista quisiera afirmar que una teoría es aproximadamente verdadera aún cuando sus términos no refieran* y agrega: si no hubiese entidades, tales como los genes, entonces la teoría genética, sin importar qué tan bien confirmada esté, no podría ser estimada como aproximadamente verdadera. Ahora bien, lo que la historia de la ciencia nos ofrece es una plétora de teorías que fueron, tanto exitosas, como no-referenciales. Estas son algunas de ellas:

- La teoría de las esferas cristalinas de la astronomía antigua y medieval.
- La teoría de los humores en medicina.
- La teoría del efluvio de la electricidad estática.
- La geología catastrofista.

- La teoría del flogisto.
- La teoría del calórico.
- Las teorías de las fuerzas vitales.
- El éter electromagnético.
- El éter óptico.
- La teoría de la inercia circular.
- Las teorías de la generación espontánea (Laudan 232-33).

El análisis anterior es la base para la defensa del argumento que Laudan ofrece contra el realismo y que es ampliamente reconocido en la literatura como 'metainducción pesimista'. Se trata de un argumento que se puede resumir de la siguiente manera:

La historia de la ciencia está llena de teorías que en diferentes tiempos y por largos períodos han sido empíricamente exitosas, pero que se revelaron como falsas en sus afirmaciones sobre la estructura íntima del mundo. De igual modo, tales teorías están llenas de términos teóricos que no refieren. Por lo tanto, por simple (meta) inducción, nuestras actuales teorías son probablemente falsas -o es más probable que sean falsas que verdaderas- y muchos o la mayoría de los términos teóricos pueden ser no-referenciales.

Por lo tanto, el éxito empírico de una teoría no proporciona una garantía a favor de la afirmación de que la teoría es aproximadamente verdadera. (Psillos 101)

A la luz de este argumento, Laudan desea rechazar la legitimidad de la conexión entre verdad aproximada y éxito. Sin embargo, Diéguez considera que el argumento se propone rebatir una tesis que el realista, no mantiene i.e., que el éxito de una teoría justifica su verdad. Un examen de esta crítica, debe indicar primero que, *sensu estricto*, aquello que Laudan atribuye al realista es la afirmación según la cual, el éxito de una teoría lo autoriza a inferir su verdad aproximada. De hecho, la inducción pesimista pretende probar que estaríamos más bien autorizados a inferir la falsedad de nuestras actuales teorías y no su verdad aproximada.

Como otra de las líneas de ataque a la inducción pesimista busca mostrar que: 'teorías falsas pueden ser estimadas legítimamente como aproximadamente verdaderas', pasará al análisis de algunas de las críticas de Psillos.

Críticas de Psillos

Para reconstruir las críticas de Psillos contra la metainducción pesimista, conviene primero indicar, la estructura del argumento atribuida por él a Laudan. Después mostraré las principales líneas de su argumentación.

Psillos considera que el argumento de Laudan pretende mostrar que la historia de la ciencia destruye la credibilidad de la explicación realista del éxito de la ciencia, sobre la base de un argumento similar a un tipo de 'reducción' y cuyo objetivo es rebatir la siguiente tesis realista:

- (A) las teorías actuales -exitosas-, son aproximadamente verdaderas.

Además, él piensa que Laudan no rechaza directamente la posibilidad de que nuestras actuales teorías sean verdaderas, sino que se propone más bien desacreditar la afirmación sobre la existencia de una conexión explicativa, entre el éxito empírico y la aproximación a la verdad, lo cual constituye la garantía de la afirmación realista. Para rechazar dicha conexión Laudan defiende las siguientes dos tesis:

- (B) si actualmente las teorías exitosas son aproximadamente verdaderas, entonces las teorías pasadas no pueden haberlo sido.

Las teorías pasadas no pueden considerarse como aproximadamente verdaderas, porque las entidades en ellas postuladas no existían o porque las leyes y los mecanismos en ellas postuladas no son parte de las actuales descripciones teóricas del mundo. Por otro lado, la segunda tesis afirma que:

- (C) estas teorías característicamente falsas fueron, sin embargo, empíricamente exitosas.

Psillos mantiene que el realista, puede conceder que todas las teorías científicas son probablemente falsas. Sin embargo, a su juicio, una teoría falsa, puede ser aproximadamente verdadera, debido a que una teoría puede ser estimada como aproximadamente verdadera de manera genuina, si describe un mundo que es similar al mundo actual en la mayor parte de sus características centrales o relevantes. Por lo tanto, lo que el realista debe mostrar es que, aunque algunas de las teorías pasadas -exitosas-, fuesen refutadas, podrían ser consideradas

como aproximadamente verdaderas. Otro de los retos del realista, consiste en poder responder al desafío de Laudan, de que no se puede mantener consistentemente la tesis de que las teorías científicas son aproximadamente verdaderas aunque sus términos centrales no refieran.

Para responder a estas objeciones y retos, Psillos combina dos estrategias. La primera consiste en intentar reducir el número de contraejemplos para mostrar que la base inductiva no es tan amplia, ni lo suficientemente representativa como garantía de la inducción pesimista. Para defender este punto muestra que hay serias dudas de que todas las teorías de Laudan pertenezcan a un estadio de desarrollo lo suficientemente maduro o que sean genuinamente exitosas. La segunda estrategia, consiste en ofrecer una noción de éxito empírico mucho más rigurosa que la de Laudan.

En desarrollo de esta estrategia, Psillos mantiene que un análisis atento al desarrollo histórico de ciertas disciplinas revela, por ejemplo, que el éter gravitacional de LeSage y Hartley, las esferas cristalinas y la teoría de la inercia circular, que aparecen en el listado de Laudan, no son genuinamente exitosas. Adicionalmente, si exigimos que superen el test de la madurez científica, tanto la teoría humoral en medicina, como la teoría del efluvio también tendrían que ser eliminadas.

Sin embargo, para hacer aún más contundente el rechazo del argumento metainductivo, Psillos se propone ofrecer una noción de éxito empírico, mucho más exigente que la de Laudan. A su juicio, una teoría es exitosa *syss*, permite hacer predicciones novedosas que sean, en principio, testables. En esta dirección, una predicción es novedosa *syss*, predice un fenómeno cuya existencia, está determinada solamente después que la teoría sugiere su existencia, i.e., si el fenómeno era hasta entonces desconocido.

Psillos advierte que se podría objetar a esta noción de éxito lo siguiente: sí sólo una predicción novedosa puede conferir soporte a una teoría, entonces podríamos admitir que una vez somos conscientes de que el hecho fue conocido a partir de la teoría, dicho hecho ya no podría soportarla más. A este reparo, Psillos responde que la predicción P de un hecho conocido es novedosa -relativa a una teoría T- si ninguna información sobre este fenómeno fue usada en la construcción de la teoría que lo predice (106) i.e., si no cumple ninguna de las siguientes condiciones de *ad hocidad*:

1. Un cuerpo de conocimiento B implica la existencia de E. La información sobre E es usada en la construcción de una teoría T y T se acomoda a E.
2. Un cuerpo de conocimiento B implica la existencia de un fenómeno E. Cierta teoría ya disponible (T) no predice, ni explica a E. T es modificada por T*, por lo que T* predice a E, pero la única razón para esta modificación es la predicción-explicación de E. En particular, T* no tiene otro exceso sobre T, ni en el contenido teórico, ni en el contenido empírico (106-07).

Sobre esta base, una predicción novedosa es aquella en la que la predicción de un fenómeno E, se hace con respecto a una teoría T, si E es desconocido antes de que se propusiera T y si no satisface ninguna de las condiciones de *ad hocidad* arriba indicadas.

Psillos sostiene que al aplicar este test, no logran sobrevivir el resto de las teorías indicadas por Laudan. No obstante, todavía quedan dos teorías, la teoría del calórico y la del éter óptico, que logran sobrevivir a dicho test y que además, son evidentemente no referenciales. Para responder a estos dos contraejemplos, Psillos emplea la estrategia denominada: '*the divide et impera move*', la cual se propone atacar principalmente la premisa (B).

Según Psillos, para rebatir (B), basta con mostrar que las leyes teóricas y los mecanismos que generaron el éxito de teorías pasadas, han sido retenidos en nuestra imagen científica actual (108). Que esto es posible, queda ilustrado, debido a que cuando una teoría es abandonada, sus constituyentes teóricos, i.e., los mecanismos teóricos y las leyes, no son rechazados en su totalidad.

Aunque Psillos concede que algunos mecanismos explicativos son inconsistentes con las teorías recientes, no todos lo son. Algunos de ellos, han sido retenidos como constituyentes esenciales de teorías subsecuentes. Por esto es necesario identificar los constituyentes de teorías pasadas, genuinamente exitosas, que hicieron contribuciones esenciales a sus sucesoras.

En defensa de esta tesis, Psillos examina la teoría del calórico, indicando cómo sus constituyentes esenciales fueron preservados en teorías subsecuentes. Adicionalmente, Psillos mantiene que el calórico no puede ser estimado legítimamente como una entidad putativa, dado que en la

teoría no eran esenciales las afirmaciones concernientes a la existencia de un tipo de fluido imponderable que causaba el aumento o el descenso de la temperatura cuando era absorbido por un cuerpo. De lo anterior se colige que sus leyes y derivaciones eran independientes de la hipótesis que afirmaba que la causa del calor era una sustancia material (113).

Ahora bien, en respuesta a estas objeciones, consideremos la primera estrategia de Psillos, i.e., aquella con la que intenta reducir la lista de Laudan, defendiendo que no se trata, en modo alguno, de teorías que pertenezcan a estadios lo suficientemente maduros de la ciencia. En lo que concierne a esta estrategia, creo que conviene indicar que Laudan examina una posible maniobra del realista para escapar a la metainducción pesimista, la cual consiste en afirmar que dicho análisis, sólo se aplica al dominio de 'las ciencias maduras'. A juicio de Laudan, esta maniobra es insatisfactoria en dos aspectos:

Primero, corre el riesgo de hacer al REC vacuo, debido a que sus defensores, definen a una ciencia madura, como aquella, cuya correspondencia o relación de casos límite se puede obtener invariablemente entre cualquier conjunto de teorías una vez han cruzado el umbral de la madurez científica. Sin embargo, semejante punto de vista hace inevitablemente tautológica la definición de 'madurez científica', porque no hay observación histórica, que pudiese concebiblemente refutar o confirmar esta tesis, debido a que aún si encontráramos una ciencia en la que se dieran las relaciones de correspondencia entre una teoría sucesora y una predecesora, no tendríamos manera de saber si tal relación, continuaría aplicándose a subsecuentes sustituciones teóricas en dicha ciencia (Laudan 233).

La segunda dificultad estriba en el propósito manifiesto del realista es el de explicar por qué la ciencia es exitosa; pero el asunto radica en que partes de la ciencia, incluidas muchas ciencias inmaduras, han sido exitosas por muchísimo tiempo. Por lo tanto, si el realista se restringe a sí mismo, a explicar solamente cómo trabaja la ciencia madura, falla en su ambición de explicar, por qué la ciencia es *en general* exitosa.

A mi juicio, Psillos pasa por alto la exigencia de Laudan, según la cual una defensa articulada del realismo debe mostrar cómo el realismo logra ofrecer una explicación general del éxito de la ciencia. De lo contrario, tendríamos que inferir la consecuencia epistemológica indeseable de que la ciencia es una empresa exitosa, tanto en estadios primitivos como en

estadios maduros. Y aunque disponemos de una explicación razonable acerca de por qué lo es en episodios maduros, ignoramos por completo, por qué también es exitosa en estadios no avanzados. Sí así fuera, ¿cuál sería la diferencia entre ciencias maduras y primitivas?

En la segunda parte de su estrategia, Psillos se propone defender que teorías falsas pueden ser estimadas como aproximadamente verdaderas, de manera genuina. Conviene indicar que la tesis defendida aquí es obviamente, la tesis de la convergencia. En su defensa, Psillos argumenta que la existencia y las propiedades atribuidas al calórico, por ejemplo, no se requerían para derivar las leyes calorimétricas³. Una parte sustancial de la estrategia de Psillos, consiste en apelar a la teoría causal de la referencia para defender que, nuestras mejores teorías recientes fijan la referencia de los términos empleados por las teorías pasadas, por ejemplo, el ‘campo electromagnético’ fijaría la referencia de la locución ‘éter luminoso’ (Doppelt 107-08).

Sin embargo, una estrategia como ésta, adolece del no despreciable defecto de que no es evidente que la teoría causal pueda explicar los fracasos en la referencia⁴. Después de todo, estamos autorizados a esperar que una teoría realista de la referencia no sólo explique el éxito de la referencia, sino también por qué fracasa.

Conclusiones

En este artículo, he examinado dos de las líneas de ataque que se han dirigido en contra del argumento metainductivo. La primera de ellas, empleada por Diéguez, afirma que Laudan atribuye ilegítimamente un conjunto de premisas ocultas (o supuestos) a los argumentos realistas. He mostrado que el primer supuesto de Diéguez se puede inferir por *Modus Tollens* de las premisas de los argumentos realistas.

³ Psillos limita el compromiso realista con la verdad, de aquellos componentes de una teoría que son requeridos para derivar sus nuevas predicciones, indicando que tales componentes son los únicos responsables de las nuevas predicciones. Sin embargo y como lo ha mostrado Doppelt, si concebimos el realismo científico, como una hipótesis filosófica, que no sólo se propone explicar el éxito predictivo de las teorías científicas, sino también, el éxito explicativo, tenemos que aquellos componentes -falsos-, que han sido eliminados, son los responsables del éxito explicativo de la teoría. ¿Cómo puede el realista explicar esto?

⁴ Como lo ha indicado Cummeskey: “los defensores de la teoría causal se han concentrado más en explicar cómo los términos de la ciencia pueden referir a los mismos aspectos de la realidad, cuando las teorías sobre aquellos aspectos de la realidad cambian radicalmente, por lo que han descuidado la explicación acerca de por qué fracasa la referencia” (Cummeskey 26).

En lo que concierne al segundo supuesto, es absurdo debilitar la afirmación realista como Diéguez sugiere, porque si así fuera, el realista, tendría que admitir que la referencia es sólo ocasionalmente, el agente causal que explicaría el éxito de la ciencia.

En el análisis del tercer supuesto, he dicho que Laudan no atribuye al realista la afirmación de que el éxito instrumental de las teorías justifica su verdad; sino más bien, su aproximación a la verdad. La fuerza del argumento metainductivo, estriba en que en lugar de disponer de evidencia histórica que nos autorice a inferir su aproximación a la verdad, tendríamos que inferir, más bien, su aproximación a la falsedad.

La segunda línea de ataque es ofrecida por Stathis Psillos, quien se propone rebatir el argumento de Laudan intentando reducir el número de contraejemplos y recurriendo finalmente, a la teoría causal, para establecer una continuidad suficiente en la referencia de los términos teóricos de las teorías científicas. Para responder a tales objeciones, he indicado que una buena parte de la línea de ataque de Psillos -anticipada por Laudan-, logra responder de manera satisfactoria y categórica a la maniobra realista que intenta reducir la lista, indicando que las teorías propuestas por Laudan, no pertenecen a estadios maduros de la ciencia. Si ello fuese así, el realista tendría que admitir que no dispone después de todo *de una explicación general del éxito de la ciencia*.

En lo que concierne a la maniobra argumentativa de Psillos, de apelar a la teoría causal de la referencia para establecer la continuidad de la referencia de los términos teóricos, he sugerido que una estrategia como ésta, enfrenta el defecto no despreciable de que no disponemos de una teoría causal satisfactoria que explique -no sólo el éxito- de la referencia, sino también su fracaso.

REFERENCIAS BIBLIOGRÁFICAS

Cummeskey, D. "Reference failure and scientific realism: a response to the Meta-induction". *British Journal Philosophy of Science*. Jan. 1992: 21-40. Oxford University Press. Print.

Diéguez, Antonio. *Realismo científico: una introducción al debate actual en la filosofía de la ciencia*. Málaga: Universidad de Málaga, 1998. Impreso.

Doppelt, G. "Reconstructing scientific realism to rebut the pessimistic Meta-induction". *Philosophy of Science*. 2007. 107-108. University of Chicago Press. Print.

Laudan, Larry. "A confutation of convergent realism". *Philosophy of Science*. 1981: 19-49. University of Chicago Press. Print.

Psillos, Stathis. *Scientific realism: how science tracks truth*. New York: Routledge, 1999. Print.



**BREVES REFLEXIONES SOBRE EL DESAFÍO DEL
ACTIVISTA, A LA POLÍTICA DELIBERATIVA: EL BUEN
DELIBERATIVISTA, NO SIEMPRE DELIBERA**

*BRIEF THOUGHTS ON THE ACTIVIST'S CHALLENGE TO DELIBERATIVE POLITICS: THE GOOD
DELIBERATIVIST DOES NOT ALWAYS DELIBERATE*

MACARENA MAREY

Universidad de Buenos Aires - Conacyt, Argentina. macarenamarey@hotmail.com

RECIBIDO EL 6 DE MARZO DE 2012 Y APROBADO EL 24 DE MAYO DE 2012

RESUMEN ABSTRACT

En este artículo, se presentan algunas reflexiones acerca de cómo los demócratas deliberativos deberían responder al "desafío del activista".

The aim of this article is to set forth some reflections on how deliberative democrats should respond to the 'activist challenge'.

PALABRAS CLAVE KEY WORDS

Deliberación, desafío del activista, legitimidad, normativa.

Deliberation, activist challenge, legitimacy, normativity.

El “desafío del activista” planteado por Iris Young, condensa una objeción frecuente contra la idea de participación ciudadana de la política deliberativa. Young construye el personaje del ciudadano comprometido con el ideal deliberativo, proponiendo que el puro intercambio racional, ordenado y no coercitivo de argumentos, realizado con el objetivo de generar consenso, sigue siendo la mejor vía de acción para promover la justicia *incluso en contextos estructuralmente injustos y excluyentes*.

El problema que plantea la crítica de Young a la democracia deliberativa se articula alrededor de dos indicaciones que, de ser correctas implicarían ciertamente que la insistencia en que la sola deliberación como la mejor vía para la resolución de conflictos y toma de decisión, *en cualquier contexto dado*, equivale a proponer mecanismos que reproducen un *statu quo* injusto. Estas dos observaciones críticas consisten en que los procedimientos deliberativos concretos, en contextos injustos son cerrados y excluyentes, porque el *ideal mismo* de la deliberación implicaría una idea meramente formal de la inclusión en la participación y necesariamente insensibles al hecho de que las sociedades contemporáneas están atravesadas por desigualdades estructurales injustas.

El núcleo más fuerte de la crítica del desafío de Young consiste, a mi juicio, en que sugiere que la consecuente insensibilidad al contexto de aplicación imputada a los procedimientos reales de deliberación estaría causada por *los presupuestos teóricos mismos de la situación deliberativa ideal*. Por este motivo, el desafío del activista llama la atención sobre la cuestión de si la distancia entre teoría deliberativa y praxis es un problema conceptual más serio que una mera cuestión técnica –la crítica de Young debe- por tanto, ser tomada en serio por los deliberativistas.

Young define a la democracia deliberativa como una teoría que ofrece una “visión normativa de las bases de la legitimidad democrática” y una “prescripción para cómo los ciudadanos deben involucrarse políticamente” (Young 672). En cuanto a esto último, para Young el deliberativismo consiste en sostener que: “la mejor y más apropiada manera de conducir la acción política, de influir y tomar decisiones públicas es por medio de la deliberación pública”.

En este artículo objetaré esta segunda tesis de Young sobre la política deliberativa, a partir, precisamente, de la afirmación de que el deliberativismo es una visión normativa y evaluativa de la democracia, por contraposición a una concepción descriptiva de ella.

El breve análisis que realizo se centra sobre todo, en el carácter normativo de la democracia y la política deliberativa y además, en la idea de que la deliberación refiere, en tal marco normativo-evaluativo, a un modo determinado de *justificación de decisiones políticas* y a un conjunto de criterios evaluativos para una crítica sólida a las relaciones de desigualdad en las sociedades políticas concretas, antes que a una mera forma de comunicación que, en virtud de su supuesta capacidad instrumental para conducir a resultados epistémicamente correctos, el bien común o la mejora de los juicios y actitudes de los ciudadanos, es considerada superior a otras formas vistas como ilegítimas por su incapacidad para ello (esto es, antes que a la imagen de la política deliberativa presupuesta por las críticas de Young).

La respuesta básica al desafío del activista que propongo consiste en que el ideal deliberativo del intercambio recíproco de razones *no implica analíticamente* la creencia de que las virtudes de los procedimientos deliberativos sean tan elevadas, que conduzcan a resultados justos (o más justos que el *statu quo* anterior a la aplicación de un procedimiento deliberativo), en contextos en los que no se verifica algún umbral mínimo de igualdad y justicia social (precondiciones para la legitimidad plena del intercambio de razones).

En otras palabras: en la medida de que el fin que guía su conducta política sea conseguir un escenario de legitimidad deliberativa, el deliberativista no tiene por qué renunciar a la ironía, las manifestaciones y otros medios estratégicos generalmente considerados disruptivos. Esta hipótesis en apariencia trivial (que virtualmente ningún deliberativista negaría) acerca de la democracia deliberativa es lo que el desafío de Young desconoce; por tanto, mostrar cuáles son las razones principales por las cuales la democracia deliberativa sostiene tal hipótesis es una tarea que, si bien humilde, no deja de ser importante para los deliberativistas.

Algunas respuestas al desafío del activista: deliberar no siempre es la respuesta

Una respuesta convincente a la crítica de Young es ofrecida por Fung, quien sostiene que es falso que del ideal demócrata deliberativo se siga necesariamente una ciega exhortación a ponerse a deliberar en cualquier contexto político, tal como postula Young.

Fung concuerda con Young en que la deliberación no es el procedimiento indicado para alcanzar resultados justos en cualquier contexto concreto dado. Consecuentemente, propone una *ética política* que sirve como guía para la acción orientada a conseguir escenarios sociales en los cuales el impacto de las desigualdades injustas sobre la deliberación pública se vea neutralizado o, por lo menos disminuido.

La argumentación de Fung comienza por indicar que la democracia deliberativa es “un ideal político revolucionario” en dos sentidos: por un lado, porque

llama a cambios fundamentales en las bases de la toma de decisión política, en el alcance de aquellos incluidos en los procedimientos de decisión, en las instituciones que hospedan esos procedimientos y, por lo tanto, en el carácter mismo de la política. (Fung 397)

Y por el otro, porque requiere

condiciones políticas, sociales y económicas dramáticamente más igualitarias que las que existen en cualquier sociedad contemporánea. (Ibíd. 397-98)

El segundo paso de la respuesta en cuestión consiste en reconocer que el mismo carácter “revolucionario” de la democracia deliberativa no autoriza al demócrata deliberativo a despachar rápidamente el desafío del activista. Fung nota acertadamente que es precisamente el hecho de que los ideales deliberativos tengan esta “naturaleza revolucionaria” lo que los vuelve “incompletos” y esto en un sentido específico:

En particular, ofrecen poca guía en lo que atañe a las responsabilidades de los demócratas deliberativos en las circunstancias decididamente no ideales que caracterizan la política democrática. Además, carecen de una teorización acerca de cómo las instituciones y prácticas existentes pueden volverse más deliberativas. Estas omisiones han llevado a algunos críticos de la democracia deliberativa a malinterpretar las afirmaciones de sus proponentes. (Fung 398)

Sin una ética de la acción política, sostiene Fung finalmente, que proponga *medidas no-deliberativas* para “mitigar exitosamente los efectos

de la desigualdad sobre la deliberación y para moderar la reticencia de las partes poderosas a actuar con reciprocidad deliberativa” (Ibíd. 312), el activista deliberativo, enfrentado con las profundas desigualdades sociales, económicas, culturales, políticas y de capacidades para deliberar, encontrará inútil que la teoría de la democracia deliberativa requiera “mucha más igualdad social de la que realmente existe” (Ibíd. 413).

Consistentemente, Fung propone una serie de principios de acción que, *sin contradecir el ideal deliberativo*, dan forma a cursos de acción *no-deliberativos en la práctica* para disminuir el impacto regresivo que la estructura profunda de las desigualdades derrama sobre los procedimientos de decisión política.

Los principios propuestos por Fung son los de fidelidad, caridad, agotamiento de instancias y proporcionalidad y actúan como guías de la práctica no-deliberativa del activista deliberativo en contextos progresivamente injustos, manteniendo siempre en vista que *el fin* (normativo) de estas medidas es conseguir que la toma de decisión se realice a partir de procedimientos plenamente deliberativos. Con esto, la práctica política no-deliberativa sigue siempre orientada a conseguir un escenario político deliberativo más justo y no, por ejemplo, al mero triunfo de un interés particular.

En un trabajo reciente, Estlund elabora una crítica a ciertas propuestas demócratas deliberativas que es similar a la crítica de Young.

Estlund apela al “teorema del segundo mejor”, desarrollado por Lipsey y Lancaster¹, para rechazar la plausibilidad de lo que llama la “doctrina del reflejo” (*mirroring doctrine*), sin que esto implique un rechazo pleno del ideal deliberativo.

¹ La falacia de la idea de que un contexto que satisface más condiciones óptimas (o más condiciones de una situación ideal), sin satisfacer todas esas condiciones, es mejor que otro contexto en el que se satisfacen menos de estas condiciones, fue criticada por Lipsey y Lancaster con el “teorema del segundo mejor”. La teoría general del segundo mejor indica que la no satisfacción de una condición óptima o ideal modifica la *deseabilidad* de las demás condiciones óptimas o ideales en ese contexto específico (por tanto, modifica la deseabilidad de la situación Pareto-óptima original): “el teorema general [del segundo mejor] afirma que si en un sistema general de equilibrio se introduce una restricción que impide la obtención de una de las condiciones óptimas, las otras condiciones óptimas dejan de ser deseables, aunque sean todavía alcanzables”. Una de las consecuencias prácticas de esto es que “dado que una de las condiciones Pareto-óptimas no puede ser satisfecha, entonces se puede alcanzar una situación óptima solamente apartándose de todas las demás condiciones Pareto-óptimas. La situación óptima finalmente alcanzada puede ser llamada un ‘óptimo segundo-mejor’ porque es alcanzada bajo una restricción que, por definición, impide que se alcance un Pareto-óptimo” (Lipsey & Lancaster 11).

Según Estlund, esta doctrina del reflejo se aplicaría al ámbito de los deberes de civilidad y de la regulación de la comunicación pública en la esfera política informal y sostendría que la participación ciudadana debe en todo momento regirse por el criterio de asemejarse en todo lo posible a la situación deliberativa modelo, sin consideración de las particularidades del contexto. Tal doctrina sostendría, por ello, una idea de participación similar a la que adopta el personaje del ciudadano deliberativo construido por Young.

El argumento central por el que Estlund rechaza la doctrina del reflejo se basa en la idea de que los deberes de civilidad deben entenderse como deberes de acción colectiva, esto es, como deberes “que se aplican solamente en la medida en que los demás los cumplen en general”.

Estlund se pregunta qué sucede “cuando la acción colectiva se quiebra”, esto es, cuando “el deber original cae”; “¿qué es lo que surge en su lugar?” (191). Su respuesta consiste en que cuando la cláusula del comportamiento cooperativo de los otros no se verifica esto no equivale a la entrada en vigencia del principio de “vale todo”. Los deberes de civilidad adquirirían, por el contrario, un “nuevo contenido sustituto”: “el deber que tenemos cambia, ajustándose a esa circunstancia” (191).

Esta idea da forma a su vez a la propuesta de Estlund de una “teoría de la ruptura” (*breakdown theory*), un “modelo de civilidad en la participación política que da un lugar a la participación brusca, disruptiva, e incluso supresora bajo las circunstancias adecuadas” (Ibíd. 185), que permite a los ciudadanos adoptar cursos de acción que *no están contemplados, según Estlund, por el diseño de la situación dialógica modelo*. El ideal deliberativo no perdería con ello, sin embargo, su vigencia normativa, debido a que

la situación deliberativa ideal, incluso existiendo solo en el pensamiento, sirve como modelo contra el cual evaluar la realidad con el fin de identificar y tratar las desviaciones. (Estlund 199)

Considero que la propuesta de una teoría de la ruptura, tal como la justifica Estlund, no ofrece buenas razones para ser adoptada por los deliberativistas como una guía de acción o ética política consistente con su objetivo político. El motivo principal de esto consiste en que la justificación normativa que Estlund ofrece para dar forma a la ética política de su teoría de la ruptura apela *exclusivamente* a sus virtudes

epistémicas, esto es, a la eficacia de tal ética política para conducir a resultados *correctos*:

Presuponiendo con Marcuse y Mill que el valor de la deliberación ordenada consiste en que promueve la verdad, o sabiduría, o calidad de las decisiones sociales resultantes, la civilidad estrecha [i. e., la doctrina del reflejo] ya no promueve la verdad una vez que están ausentes los demás componentes de una deliberación ordenada pero libre y cuando los criterios que permiten desviaciones de la civilidad estrecha pueden servir para remediar la situación epistémica. (Estlund 193)

Estlund sostiene, por ende, que cuando la cláusula del comportamiento cooperativo de los demás no se verifica, la obligación de civilidad que permanece sigue manteniendo la misma fuente de normatividad que daba forma al deber de civilidad original. Esta fuente de normatividad es, para Estlund, el *fin epistémico de alcanzar decisiones correctas*. Como es sabido, la propuesta demócrata de Estlund, en su conjunto, es epistémica y, por ello, su valoración de la deliberación depende eminente y necesariamente de su capacidad instrumental para producir resultados que muestren una calidad epistémica superior a los resultados alcanzados con otros procedimientos.

Asimismo, este fin justifica normativamente no solo todos los deberes de civilidad (*se cumpla o no se cumpla* la cláusula del comportamiento cooperativo de los demás), sino que también da forma a una idea de participación determinada: se participa *para alcanzar resultados epistémicamente correctos*. Es decir: arribar a resultados verdaderos o correctos es el (único) fin que da sentido a la participación política. Con esto, aparece una tensión en la postura de Estlund por la cual se pone en relieve un inconveniente de los modelos exclusivamente epistémicos de justificación de la deliberación, inconveniente que ulteriormente implicará que esta propuesta no sea plenamente compatible con un ideal demócrata deliberativo.

Esta tensión es causada por la coexistencia de una justificación epistémica de la democracia con el énfasis puesto por Estlund en la dependencia del contexto al momento de dar forma a la participación política. El inconveniente consiste, en que cuando la dependencia del contexto es tomada como el factor que determina el tipo de participación que resulta adecuada en vistas a la consecución de resultados epistémicamente

correctos, una justificación *epistémica* de la democracia debería *dejar indeterminado cuál es el procedimiento de decisión colectiva que resulta superior a otros modelos disponibles*.

Si seguimos la argumentación de Estlund y tomamos en serio la idea de que es el contexto concreto el que determina de manera suficiente siempre el tipo de procedimiento de decisión, no tendríamos manera de justificar la preferibilidad de un tipo específico de legitimidad procedimental democrática *ex ante* la aplicación de procedimientos concretos, en contextos concretos. Podría ocurrir que en un contexto concreto dado la mera agregación de preferencias individuales sea epistémicamente superior a la deliberación comunal acerca del bien común. En consecuencia, el esquema de argumentación de Estlund no es suficiente para mostrar por qué deberíamos orientar la praxis política hacia un ideal deliberativo, ni por qué deberíamos deliberar políticamente en general en primer lugar, dado que la capacidad instrumental de cualquier procedimiento de decisión colectiva para producir resultados correctos es una virtud que, según se infiere del tratamiento de Estlund de la participación ciudadana, se verifica o no dependiendo en todos los casos del contexto en el que se aplica.

Por este motivo, la propuesta de Estlund resulta finalmente inconsistente con un ideal deliberativo, e incluso resulta inconsistente con mantener el ideal deliberativo como un criterio regulativo contra el cual evaluar la realidad política concreta. Consecuentemente, la propuesta epistémica de Estlund no consigue resguardar (a pesar de que se lo propone) la vigencia del estatuto normativo del modelo de legitimidad deliberativa y este motivo es, en mi opinión, suficiente para considerar que su teoría de la ruptura no es una guía de la participación política plenamente consistente con el deliberativismo.

El carácter normativo de la democracia deliberativa y algunos malentendidos

La propuesta del activista deliberativo de Fung tiene algo en común con la objeción de Young -un punto al que también apela Estlund-. En efecto, ambas se derivan, en última instancia del hecho de que la democracia deliberativa es un tipo de teoría *normativa* de la democracia. Por un lado, la objeción del activista se concentra en la *incapacidad* de las teorías deliberativas para franjear la distancia entre teoría normativa y

praxis; por su parte, la respuesta de Fung consiste en proponer que la teoría de la democracia deliberativa no tiene por qué adoptar una idea de la participación política que dificulte innecesaria y ciegamente la concreción del modelo ideal normativo y que esto se debe a las exigencias normativas que hace la teoría misma.

Este rasgo de la política deliberativa (i.e., que ella sea una teoría normativa de la democracia) nos permite comenzar a responder las dos objeciones de Young. En lo que sigue, señalaré (con las letras A-D) los pasos de esta respuesta en apariencia trivial al desafío del activista. A partir del análisis de tres problemas que Sanders plantea como objeciones que refutarían el *ideal teórico* de la democracia deliberativa², quisiera ejemplificar, así mismo, cómo los desafíos a la noción de participación de la democracia deliberativa pueden ser superados apelando a su carácter normativo.

Young no negaría que A) dado su carácter normativo, la política deliberativa puede ofrecernos criterios evaluativos que nos permiten explicar por qué un contexto político determinado es estructuralmente injusto y excluyente y por qué el hecho de que un contexto tenga tal característica pone en cuestión la justicia y la legitimidad de los procedimientos de razonamiento práctico colectivo (y de sus resultados) aplicados en él. Con esto, el ideal deliberativo se convierte también en una *teoría evaluativa de la democracia*.

De este modo, *prima facie*, la idea de la política deliberativa nos facilita principios gracias a los cuales podemos ejercer una crítica sólida y fundamentada contra las mismas estructuras sociales concretas que hacen que la instrumentación apresurada o prematura de esquemas deliberativos en ellas sea contra-productente.

Ahora bien, B) sostener que el ideal deliberativo muestra esta capacidad de crítica es diferente –y esto es lo que la crítica de Young supone que el *deliberativista no reconoce*– de adscribir a la deliberación una capacidad para solucionar, por sí misma, las desigualdades estructurales perniciosas.

El punto aquí es que la equiparación entre estas dos tesis es lo que da forma a la imagen que Young supone de la democracia deliberativa: lo que el *desafío del activista atribuye al deliberativista es, en última instancia,*

² Me concentraré en tres de los problemas que señala Sanders. Podemos llamarlos: “desigualdad de autoridad epistémica”, “dinámica de grupo viciada” y “atención preferente a lo común”.

esta equiparación incorrecta. Dado que es evidente que sostener un ideal deliberativo no nos compromete con sostener la tesis *pragmática* de que deliberar en cualquier contexto práctico concreto sea la mejor vía para promover la justicia social (justicia que es a su vez ciertamente precondition del correcto desenvolvimiento de un foro *plenamente* deliberativo), C) podemos adherir a la política deliberativa sin tener que negar que la sola deliberación en contextos concretos resulte incapaz de solucionar muchos de los problemas que atraviesan nuestras sociedades impidiendo que se verifiquen ciertas condiciones necesarias para la legitimidad política plenamente deliberativa.

Es, simplemente, este motivo evidente lo que debe estructurar la repuesta deliberativista al desafío de Young; en otros términos, esta es la razón por la cual es sencillamente falso que el ideal deliberativo demande analíticamente ponerse a intercambiar argumentos de manera ordenada en cualquier contexto dado, tal como supone Young.

De este modo, en lo que atañe al rol de las críticas anti-deliberativistas instadas por los estudios acerca de las “patologías deliberativas”, estoy de acuerdo con J. Cohen en que estos estudios

deben ser tratados con cierto cuidado. Aquellas patologías pueden aparecer en la decisión grupal conducida sin esfuerzos por evitar los resultados patológicos. Tales estudios pueden ser interpretados en su mayoría como fuentes de notas de precaución y recomendaciones para la mejora de la deliberación antes que como argumentos que socavan su defensa. (Cohen 339)

En su artículo “Against deliberation”, Sanders enumera y analiza algunas de estas patologías, con el fin de socavar la defensa no solamente de la aplicación de procedimientos deliberativos en contextos injustos, sino también del *ideal teórico* mismo de la política deliberativa.

Sin embargo, según lo dicho anteriormente, podemos aceptar que la mera aplicación de mecanismos de intercambios ordenados de argumentos no es suficiente para neutralizar tales patologías, sin tener que rechazar, por ello, el mencionado ideal. Un somero análisis del modo en que Sanders considera que las patologías de la discusión grupal afectan a la deliberación revela que ella se maneja con una imagen de la política deliberativa que tiende, precisamente, a confundir los dos

aspectos que he diferenciado, i.e., el estatuto normativo y evaluativo de la democracia deliberativa, por un lado, y su supuesta capacidad instrumental para neutralizar los inconvenientes que distorsionan el correcto desenvolvimiento de los foros públicos políticos concretos, por el otro.

En primer lugar, Sanders nota que

la deliberación requiere no solamente igualdad de recursos y la garantía de igual oportunidad para articular argumentos persuasivos, sino también igualdad de 'autoridad epistemológica', de capacidad para evocar reconocimiento de los argumentos propios. (349)

En segundo lugar,

algunas personas pueden ser ignoradas sin importar cuán buenas sean sus razones, no importa cuán hábilmente las articulen, y cuando esto sucede, la teoría democrática no tiene una respuesta porque no se puede contrarrestar una dinámica de grupo viciada con una buena razón. (Sanders 354)

La consecuencia que de esto extrae Sanders para la teoría de la democracia deliberativa consiste, en que

Incluso si los teóricos democráticos reconocen las inequidades asociadas con clase, raza, género y recomiendan, por ejemplo, igualar el ingreso y la educación para redistribuir los recursos necesarios para la deliberación –incluso si todos pueden deliberar y aprender a dar razones– las ideas de algunas personas pueden seguir contando más que las de las otras. Insidiosos prejuicios pueden inclinar a los ciudadanos a escuchar algunos argumentos y no otros. Importantemente, este prejuicio puede no ser reconocido por aquellos ciudadanos cuyas visiones son ignoradas. (Ibíd. 353)

Estas dos indicaciones, sumadas, implican que aplicar procedimientos deliberativos en contextos viciados en los que no se verifica algún nivel mínimo de igualdad de "autoridad epistemológica" no redunde necesariamente en mecanismos de decisión plenamente inclusivos.

Sin embargo, podríamos responder a Sanders que los procedimientos reales de intercambio de razones no serán, en tales contextos, fieles a las exigencias normativas que hace la teoría de la política deliberativa; es decir, que no podrán ser llamados plenamente deliberativos precisamente porque en el contexto de aplicación no se verifica la preconditione necesaria de igualdad en la capacidad para evocar “autoridad epistemológica”.

Ahora bien, si el ideal deliberativo puede dar cuenta de ese modo de la ilegitimidad de tales procedimientos, se puede sostener que esto se debe a que su noción de inclusión en la participación real no es meramente formal, como supone la crítica de Young. Ningún deliberativista negaría que, en palabras de Sunstein,

el valor de la deliberación, como fenómeno social, depende en gran medida del contexto social –de la naturaleza del proceso y de la naturaleza de los participantes. (124)

Todo procedimiento concreto de decisión colectiva depende de estos factores y la política deliberativa no pretende ser insensible a ellos, al modo en que son insensibles a ellos la “doctrina del reflejo” criticada por Estlund y el personaje deliberativo de Young. Si tenemos en cuenta que la justicia de todo procedimiento de decisión colectiva depende también del contexto concreto de aplicación, podemos reconocer que proponer procedimientos deliberativos cuando no se verifique algún umbral mínimo de condiciones para ello, implicaría que consideramos que la política deliberativa no es un modelo regulativo y evaluativo, sino algo diferente, a saber: la mejor *estrategia* para neutralizar los efectos negativos del hecho de que las condiciones para la deliberación no estén satisfechas por debajo de algún determinado umbral mínimo (lo cual sería poner el carro delante del caballo).

Pero dado que podemos rechazar esta última tesis acerca de la *capacidad instrumental* de la deliberación (tesis que el desafío del activista y Sanders atribuyen ciertamente al deliberativismo y que virtualmente ningún deliberativista defendería) sin rechazar el ideal deliberativo *tout court*, D) podemos acordar con Fung y siguiendo los mismos criterios normativos del ideal deliberativo, propondríamos cursos *alternativos a la deliberación* (entendida aquí como mera forma de comunicación), para mitigar el impacto pernicioso de la desigualdad respecto de, *in casu*, quiénes son escuchados.

En tercer lugar, encontramos un problema relacionado con el requisito deliberativo de respeto o reconocimiento mutuo. Sanders indica que la mayoría de las veces tal reconocimiento puede verse impedido por el hecho de que

lo que es reconocido por el que escucha es solamente lo que él puede incorporar, lo que es identificablemente similar. Mientras que lo diferente, distintivo, único o poco común puede ser articulado, [...] no recibe atención o no es reconocido. La atención preferente a lo que es común aumenta los riesgos de rechazo pleno de las perspectivas de las minorías. (361)

La pertinencia de la objeción de la “atención preferente a lo que es común” consiste en que el intercambio público de razones no redundará inmediata e inexorablemente en una mejora de la capacidad de los ciudadanos deliberantes para atender sinceramente las razones que, si bien correctamente articuladas en el foro público, son sostenidas desde posiciones sociales y culturales con frecuencia profundamente diferentes de las visiones mayoritarias y socialmente marginalizadas por estas últimas.

Claramente, esta objeción acerca de la desatención de las voces de las minorías no afecta exclusiva, ni eminentemente, a la democracia deliberativa, sino que afecta a todas las concepciones de la democracia, porque es suscitada por el reconocimiento de un hecho social virtualmente ubicuo. Es decir: todas las teorías de la decisión colectiva que postulan como ideal alguna noción de igualdad y autonomía para los participantes de la toma de decisión tienen que ofrecer alguna clase de respuesta a la indicación de que la igualdad y libertad meramente formales no son sensibles al hecho de que la decisión colectiva puede vulnerar no solamente los derechos y los intereses, sino también *las voces* de las minorías. La diferencia, por caso, entre la democracia agregativa y la deliberativa respecto de este punto consiste en que mientras que la primera oculta el hecho de que las minorías no son escuchadas detrás de la idea de que los ciudadanos actúan políticamente siguiendo solamente intereses particulares irreconciliables por definición, la política deliberativa hace foco sobre la necesidad de la inclusión de sus voces.

A su vez, la inclusión plena de las voces de las minorías no queda justificada instrumentalmente, por ejemplo, por referencia a la búsqueda

(epistémica) de qué es el bien común, sino que esta inclusión es, desde la perspectiva de la reciprocidad deliberativa demandada por un principio normativo: la política deliberativa se fundamenta también en la idea de que la democracia implica la participación efectiva de *todos los afectados por las normas* en su creación.

Con todo, la crítica más fuerte que está por detrás de la indicación de la atención preferente a lo que es común consiste en que es un tipo de distorsión de la comunicación que hace que la idea del “poder del mejor argumento” se convierta en un hierro de madera:

Si la dominación en las discusiones entre grupos no es atribuible a habilidades superiores en cuanto al razonamiento, argumentación o deliberación por parte del grupo dominante, entonces no es probable que distribuir las habilidades para la deliberación de manera más amplia vaya a solucionar los problemas de participación o influencia desiguales. En vez de esto, *mejorar la discusión democrática parece requerir intervenciones en la estructura de las deliberaciones grupales*³. (Sanders 366)

Sin embargo, esto último es, de hecho, una de las medidas que propone Fung, para quien es permisible para el demócrata deliberativo diseñar mecanismos de intervención, como por ejemplo, designar reguladores para el desarrollo de las discusiones con la intención de impedir que un individuo o grupo de individuos monopolice la dinámica del diálogo. Siguiendo a Fung -y esto es el sentido del último paso D) de la primera parte de la respuesta al desafío del activista-, podemos decir que no hay nada en el ideal de la política deliberativa que nos prohíba tomar medidas no-deliberativas para conseguir que los procedimientos concretos (en contextos injustos y excluyentes) se estructuren y desarrollen de manera tal que los “problemas de dominación y jerarquía” que obstaculizan el fin de conseguir escenarios plenamente deliberativos se vean neutralizados.

La razón de esto es (nuevamente) que suscribir al ideal de la política deliberativa no implica necesariamente creer que la deliberación goza de una capacidad instrumental superior a otros modos de actuar políticamente para neutralizar el efecto de tales problemas en la toma de decisión y para resolverlos por sí misma. Esta relación no es necesaria, a su vez, porque la justificación normativa de la política deliberativa no

³ Cursivas resaltadas por el autor.

se desarrolla en los términos de esta hipótesis pragmática, sino en los términos de la exigencia de reciprocidad en la justificación pública de las razones para votar o actuar políticamente.

Por este motivo, el deliberativista puede responder a Sanders que la existencia de los tres problemas de la desigualdad de autoridad epistemológica, las dinámicas de grupo viciadas y la atención preferente a lo que es común significan, en realidad, que no están dadas las condiciones de reciprocidad necesarias para que las deliberaciones grupales sean estructuralmente justas. Si aplicamos, por ende, los criterios del ideal normativo deliberativo para evaluar el contexto en el que aparecen estos problemas, diremos que las discusiones grupales a las que refiere Sanders no serían, en rigor, deliberativas en primer lugar. En consecuencia, podemos decir que *el ideal deliberativo ya implica conceptualmente una intervención en las discusiones grupales*: si podemos hablar de un procedimiento real plenamente deliberativo, es porque la influencia de estos problemas en él han sido neutralizadas.

Así mismo, podemos volver a señalar que estos problemas afectan a toda concepción de la democracia. Si queremos evaluar el potencial de la política deliberativa para operar cambios sociales, lo que debemos hacer es comparar de qué modo las diferentes concepciones de la democracia responden a los inconvenientes que plantea la incidencia de las distorsiones causadas por las desigualdades estructurales injustas sobre los juicios de los votantes. Frente a ellos, una propuesta conocida ha sido la respuesta conservadora de la línea schumpeteriana de dejar la política en manos de una elite de expertos. A diferencia de esta reacción, la democracia deliberativa apunta en la dirección contraria: para ella, la política debe permanecer en manos de los ciudadanos, esto es, *de los afectados por las normas jurídicas y las decisiones políticas en general* y las soluciones a los inconvenientes planteados por los críticos no pueden proponerse a costa de esta convicción.

Podemos, además notar que la propuesta de interrupción del activista de Young y la estrategia política del testimonio hecha por Sanders pueden ser igualmente asumidas por grupos cuyos idearios son directamente contrarios a los principios normativos de la democracia. Por tanto, podemos plantear la objeción de insensibilidad a la incidencia de las distorsiones causadas por las desigualdades injustas, también a las propuestas anti-deliberativas de Young y Sanders.

En su crítica al desafío del activista, Talisse nota acertadamente que,

aunque las discusiones de Young asocian al activista siempre con causas políticamente progresivas [...] no todos los activistas son progresistas en este sentido. Los activistas de extrema derecha racista también afirman estar peleando por la justicia, la equidad y la liberación. Sostienen que los procesos e instituciones existentes son ideológicamente hegemónicos y distorsivos. Consistentemente, rechazan cualquier ideal deliberativo por las mismas razones que el activista de Young.

[...] Lo que es más importante, consideran al vocabulario de 'inclusión', 'desigualdad estructural', 'poder institucionalizado' del activista de Young como completamente en línea con lo que ellos afirman que es una ideología hegemónica que actualmente domina y distorsiona sistemáticamente nuestros discursos políticos. (Talisse 436-7)

Por su parte, Sanders propone al testimonio como el modo de comunicación más eficaz para conseguir que las minorías hagan escuchar sus voces. Sin afirmar que el discurso testimonial deba ser excluido de los foros públicos, considero que con su propuesta aparece, en todo caso, la cuestión de cómo hemos de evaluar el contenido de un testimonio. También los ciudadanos que adhieren a ideologías anti-igualitarias son capaces de articular testimonios con la intención de generar empatía para sus posturas antidemocráticas. Asimismo, la narración testimonial estará, muy probablemente, sujeta a factores externos distorsivos, como por ejemplo su inevitable publicación en medios de comunicación masiva que cuentan con sus propias agendas y que estarán influidos por la injerencia de las desigualdades estructurales injustas.

La cuestión normativa, por ende, resurge: ¿por qué y cuándo el contenido de un testimonio determinado debería ser atendido como generador de una demanda social por parte de un colectivo minoritario injustamente excluido y por qué y cuándo otro testimonio determinado constituye en realidad una traducción de motivos anti-democráticos al estilo narrativo o emotivo?

A mi juicio, el problema de la crítica y de la propuesta de Sanders es que su concepción de la participación política se agota en la participación

misma. En efecto, Sanders desconecta la participación de su orientación inevitable hacia la toma de decisión política:

lo que es fundamental de dar testimonio es contar la historia propia, no buscar el diálogo comunal. [...] No hay una presuposición en el testimonio acerca de buscar un fin común, no hay una expectativa de discusión orientada a la resolución de un problema comunitario. (372)

Sanders sugiere con esto que la exclusión es un problema privado, que no afectaría a la comunidad en su totalidad. Desde una perspectiva democrática, la discriminación es, por el contrario, un problema que afecta no solamente al discriminado; una sociedad excluyente es una sociedad injusta en conjunto.

Por tanto, si el testimonio no se orienta a resolver un problema público y político (“comunitario”), entonces no se entiende en qué consiste su aporte social progresivo ni cómo queda justificado por referencia a un fin externo a sí mismo –con la excepción, claramente, del valioso efecto terapéutico del testimonio para la persona que lo narra. Si no se orienta a un *fin colectivo externo*, entonces “contar la propia historia” no tiene ningún efecto en la toma de decisión política. Así, la participación propuesta por el testimonio de Sanders termina o bien retirándose al ámbito privado (despolitizándose), o bien deslizándose hacia la inversión de los medios estratégicos por los fines o principios normativos de la democracia.

El objetivo *político-democrático* del testimonio no debería agotarse en contar la propia historia, sino que debería orientarse a cambiar el estado de opresión que sufre quien ofrece el testimonio. El valor público del testimonio radica en que, *por medio de él*, pueda introducirse su contenido (aquello sobre lo que él versa) en la agenda de la “discusión orientada a la resolución de un problema comunitario”.

La postura del activista entraña estos dos riesgos diferentes: la posibilidad de despolitizar la participación y la de convertirla en un fin en sí mismo desconectado de criterios normativos democráticos. Las breves reflexiones presentadas aquí sugieren que la participación política deliberativa puede ser eficaz para conseguir mayor equidad social en contextos injustos, sin por ello recaer en los inconvenientes de una concepción meramente descriptiva (no normativa, ni evaluativa), de la acción política.

Consideraciones finales

A partir de lo anteriormente visto, podemos decir que las críticas a la democracia deliberativa de la línea de Young no deben perder de vista -para no perder su eficacia- que la participación política es un medio orientado a un fin político democrático, no un fin en sí mismo incapaz de ser evaluado y criticado a partir de criterios externos de legitimidad y justicia.

Tales criterios de legitimidad no significan, sin embargo, que debamos actuar como si la utopía ya se hubiese realizado; pero tampoco la vigencia de su validez normativa y evaluativa en la práctica política de nuestras sociedades imperfectamente democráticas depende del éxito fáctico de un modo de proceder adecuado solamente en una situación ideal que -por definición- no se verifica en ninguna práctica concreta. Coincido con el análisis de Elster del tópico de la participación y de su relación con los fines políticos e ideales normativos. Criticando las dos posturas extremas (la actitud de actuar políticamente como si ya nos encontráramos en la situación ideal, por un lado, y la conversión de la participación política en un fin en sí mismo), por el otro.

hay una fuerte tensión entre dos modos de concebir la relación entre fines y medios políticos. Por un lado, los medios deberían provenir de la naturaleza de los fines, dado que de otro modo el uso de medios no adecuados podría tender a corromper el fin. Por el otro lado, existen peligros en elegir medios inmediatamente derivados del fin a ser realizado, dado que en una situación no ideal estos pueden apartarnos del fin antes que acercarnos a él. (18-9)

En los escenarios no ideales de la política concreta, el ideal deliberativo mantiene su validez normativa sin que ello implique que debamos ponernos a dialogar ordenadamente en cualquier situación dada. En términos de Estlund, diríamos que el deber de intentar conseguir un escenario deliberativo inclusivo y equitativo no se quiebra, sino que lo que cambia es el curso de acción estratégicamente adecuado y moralmente necesario para conducir a tal fin. Este fin político hace que en ciertas circunstancias no recurrir a conductas propiamente deliberativas (propias de la situación dialógica ideal) sea también un deber de civilidad. De modo que podemos coincidir nuevamente con

Elster en que el teorema del segundo mejor se aplica también a la ética política y que en este terreno no implica “la idea familiar de que algunas obligaciones morales pueden ser suspendidas cuando otras personas actúan no-moralmente”, sino que

cuando los otros actúan no-moralmente, puede existir una obligación de desviarse no solo de lo que ellos hacen, sino también del comportamiento que habría sido óptimo si hubiera sido adoptado por todos.

[...] Cosas tales como la ironía, la elocuencia o la propaganda pueden ser necesarias, implicando menos respeto por el interlocutor que lo que prevalecería en la situación discursiva ideal. (Elster 18)

El modo en que es factible mantener estos medios en el cauce de la legitimidad deliberativa democrática es asegurándonos que sea la vigencia normativa del ideal deliberativo lo que dé forma a nuestros deberes de civilidad en la participación política y no, por ejemplo, el mero interés de imponer nuestra visión de la cosa pública o nuestros intereses particulares egoístas. Finalmente, incluso al conceder que no todas las objeciones a las teorías deliberativas son falsas y carentes de pertinencia descubrimos que puede quedar un potencial crítico en la idea de la política deliberativa que no se ve realmente neutralizado por el hecho de que las sociedades se vean atravesadas estructuralmente por los problemas que instan esas objeciones. Tal potencial crítico reside en el estatuto normativo y evaluativo de la democracia deliberativa.

Resulta llamativo, por tal motivo, que Young proponga que la política deliberativa deba comenzar a responder al desafío del activista reconociendo que

la teoría democrática, incluida la teoría de la democracia deliberativa, debería entenderse a sí misma primariamente como una teoría crítica que expone las exclusiones y límites de los procedimientos concretos de decisión supuestamente justos que hacen que la legitimidad de sus resultados sea sospechosa. (Young 688)

Esta propuesta de Young llama la atención porque es, *prima facie*, contradictoria con la propia objeción que ella estaría destinada a superar y es contradictoria porque plantea la necesidad de mantener

el estatuto evaluativo (por tanto, también el normativo) de una teoría, a la vez que el mismo desafío del activista es un tipo de objeción que se ubica, precisamente, en el ámbito de la cuestión (central de la filosofía política) de cuál es el modo en que un modelo político normativo puede dar cuenta de la relación entre teoría y praxis. Frente a las dos imputaciones de Young y frente a lo que corre por detrás de ellas (que la teoría deliberativa de la democracia es incapaz de guiar la praxis en nuestras sociedades no-ideales), el desafío se relaciona directa y fundamentalmente con el modo en que se conciba la participación política ciudadana, pero también con cuál pensemos que es la tarea de la política en general. Las consideraciones básicas expuestas aquí forman parte de la razón por la cual considero que el ideal de la política deliberativa exige una idea de participación y un concepto de capacidad de actuar políticamente que son mucho más amplios que los presupuestos por la imagen que la crítica de Young mantiene de la política deliberativa. Aunque este sería tema para un trabajo ulterior, pienso que esa concepción amplia y activista de la capacidad de actuar políticamente debería ubicarse en un momento teórico *anterior* a la reflexión sobre los cursos de acción política más adecuados en los contextos concretos; esto es: *una idea amplia de participación y de acción política debería formar parte del ideal deliberativo en primer lugar y no entrar en consideración una vez que se está frente a los contextos concretos, que son no-ideales por definición.*

REFERENCIAS BIBLIOGRÁFICAS

Cohen, Joshua. "Reflections on deliberative democracy". *Philosophy, Politics and Democracy. Selected Essays*. Cambridge and London: Harvard University Press, 2009. Print.

Elster, Jon. "The market and the forum: Three varieties of political theory". *Deliberative democracy. Essays on reason and politics*. Eds. James Bohman and William Rehg. Cambridge, Mass. And London: The MIT Press, 1997. Print.

Estlund, David. "The real speech situation". *Democratic authority. A philosophical framework*. Princeton and Oxford: Princeton University Press, 2008. Print.

Fung, Archon. "Deliberation before the revolution. Toward an ethics of deliberative democracy". *Political Theory*. 2005: 397-419. Print.

Lipsey, R. G. and Kelvin Lancaster. "The general theory of the second best". *The Review of Economic Studies*. 1956-7: 11-32. Print.

Sanders, Lynn. "Against deliberation". *Political Theory*. 1997: 347-76. Print.

Sunstein, Cass. "On a danger of deliberative democracy". *Daedalus*. 2002: 120-4. Print.

Talisse, Robert. "Deliberativist responses to activist challenges. A continuation of Young's dialectic". *Philosophy and Social Criticism*. 2005: 423-44. Print.

Young, Iris Marion. "Activist challenges to deliberative democracy". *Political Theory*. 2001: 670-90. Print.

DEL CRÁTILLO AL TRACTATUS, OBSERVACIONES A UNA TEORÍA FIGURATIVA DEL SIGNIFICADO

FROM CRATYLUS TO TRACTATUS, SOME REMARKS ON PICTURE THEORY OF MEANING

ÁLVARO REVOLLEDO NOVOA

UNMSM-Universidad Nacional Mayor de San Marcos, Perú. alvarorevellido@gmail.com

RECIBIDO EL 15 DE FEBRERO DE 2012 Y APROBADO EL 18 DE MAYO DE 2012

RESUMEN ABSTRACT

En el presente artículo, presentaré el problema relativo de cómo el lenguaje, puede dar cuenta del mundo en términos representacionales desde una teoría figurativa del significado. El giro de la discusión, estará basado en algunos de los problemas presentados en el diálogo platónico el *Crátilo*, considerando dos aspectos centrales: i) cómo es, que aparentemente del conocimiento de la naturaleza del lenguaje, saltamos al conocimiento de la naturaleza de la realidad y ii) hasta qué punto el lenguaje, no es más bien un obstáculo (a la manera de un disfraz) que oculta a la realidad en vez de ser el vehículo que nos permite conocerla. No obstante, el concepto o la presunción objeto de este análisis será la idea que sostiene que el lenguaje pinta o figura los objetos del mundo. Este rasgo figurativo representacional del lenguaje, permitirá establecer el vínculo entre dicho diálogo platónico y el *Tractatus logico philosophicus* de Ludwig Wittgenstein, sumado al aspecto que considera que existen formas (la forma de los nombres o la forma lógica), las cuales permiten la correspondencia entre el lenguaje y el mundo, a través de sus elementos primitivos.

In this paper I shall set forth the relative problem of how language can provide an account of the world in representational terms from a picture theory of meaning. The main point of discussion will be based on some of the problems put forward by Plato's dialogue *Cratylus*, taking into account two central aspects: i) how is that we go, apparently, from the knowledge of the nature of language to the knowledge of the nature of reality, and ii) until what point language is not an obstacle (like a costume) that disguises reality, instead of being a vehicle that allows us to know it. However, the concept or presumption that guides this analysis is the idea that holds that language pictures or depicts the objects of the world. This figurative and representational feature of language will let us establish a link between this Platonic dialogue and Wittgenstein's *Tractatus logico philosophicus*, in addition to the view that considers that there are forms (the form of names or logical forms) that allow the correspondence between language and world through its primitive elements.

PALABRAS CLAVE KEY WORDS

Crátilo, forma lógica, lenguaje, mundo, teoría figurativa.

Cratylus, logic form, language, world, picture theory.

Introducción

Una dificultad clásica en la historia de la filosofía occidental relativa al lenguaje, es la de qué podemos hacer con él, en el plano del conocimiento, para dar cuenta del mundo. Esto supone, cuando menos, dos respuestas: la primera es positiva y sostiene que en efecto es posible conocer el mundo de las cosas, desde el lenguaje, imaginándonos algún tipo de conexión especial entre los dos, como algo que comparten a nivel estructural y que hace posible dicha relación; en cambio, si la respuesta es negativa, es decir, que el lenguaje no puede ayudarnos a conocer el mundo, la salida principal, ha sido la de sugerir alguna modalidad alternativa en la cual se trata de dar cuenta del mundo más allá de las palabras. Desde luego, existen otras posibles respuestas; tal vez, una tercera y no poco conocida es la que sencillamente defendería que el lenguaje, sea cual fuere, no tiene la capacidad de dar cuenta del mundo, razón por la cual incluso, la misma empresa de conocer la realidad, debería ser abandonada.

He querido empezar esta breve exposición, señalando algunas de las consecuencias que se desprenden toda vez que nos enfrentamos al problema del conocimiento de la realidad desde el lenguaje, sobre todo, para dejar en claro, que no es un asunto de poca monta, sino que al contrario, conforma parte de nuestra agenda de problemas filosóficos que nos impelen, directa o indirectamente, especialmente, cuando se está comprometido con estos asuntos de una manera seria. Además, mi intención, es la de clarificar, que es un problema que difícilmente podría enfrentarse si no se realiza un examen largo y detenido, involucrándose con las preguntas y el retorno de nuestros interlocutores. Por esta razón, me concentraré únicamente, en una de las respuestas sugeridas, vale la pena decir, la primera y tan solo, de manera complementaria, a las otras dos.

He titulado este artículo: “Del *Crátilo* al *Tractatus*” y ello, puede sugerir que la empresa aquí asumida es bastante pretenciosa. La verdad, es que sí lo es, sin embargo, para defender la intuición que sostengo, el subtítulo del mismo: “observaciones a una teoría figurativa del significado”, puede resultar relativamente clarificador, debido a que mi objetivo al unir estas dos obras filosóficas, es la de discutir el problema relativo, a cómo el lenguaje, puede dar cuenta del mundo en términos representacionales desde una teoría figurativa del significado. En primer lugar, el giro de la discusión, estará centrado en algunos de los problemas que presenta

el diálogo platónico el *Crátilo*, considerando dos aspectos centrales: i) cómo es, que aparentemente, del conocimiento de la naturaleza del lenguaje, saltamos al conocimiento de la naturaleza de la realidad y ii) hasta qué punto el lenguaje, no es más bien un obstáculo (a la manera de un disfraz), que oculta la realidad en vez de ser el vehículo, que nos permite conocerla. Sin embargo, el concepto o la presunción, que será objeto de análisis, es la idea que sostiene que el lenguaje pinta o figura los objetos del mundo. Este rasgo figurativo representacional del lenguaje, permitirá establecer el vínculo entre dicho diálogo platónico y el *Tractatus Logico Philosophicus*, de Ludwig Wittgenstein, sumado al aspecto que considera que hay formas (la forma de los nombres o la forma lógica), que permiten la correspondencia, entre el lenguaje y el mundo, a través de sus elementos primitivos. La observación principal a este esquema figurativo, es la consideración del significado como un asunto que resulta de la sola correlación entre el lenguaje y el mundo, lo que hace pensar, que el significado es algo fijo de una vez y no algo que se fija socialmente, como una entidad histórica.

Consecuencias del *Crátilo*

La lectura del *Crátilo*, se presenta como un diálogo en el cual sus tres interlocutores: Hermógenes, Crátilo y Sócrates, discuten un mismo problema que presentará dos tesis en conflicto: ¿los nombres son correctos, respecto de las cosas que nombran de manera natural o la corrección de los nombres, dependerá de la convención y la costumbre de quienes los usan? En esta disputa, la tesis naturalista con la que empieza el diálogo, es defendida por Crátilo quien afirmará, que independientemente del idioma, hay una rectitud natural de las denominaciones, siendo la misma para todos. La posición de Hermógenes, sostiene el punto de vista según el cual los nombres se fijan según el uso y la costumbre, y que allí radica su rectitud. Aparentemente, ambos puntos de vista son insalvables. Donde se afirma uno, se niega el otro y viceversa.

La afirmación con la que a modo de queja, Hermógenes, invita a Sócrates a participar en la discusión y en la cual expresa la posición de Crátilo, se presenta de la siguiente manera:

Hermógenes: Crátilo afirma, Sócrates, que existe por naturaleza una rectitud de la denominación para cada una de las cosas, y que ésta no es una denominación que algunos

dan –una vez han acordado darla, aplicando un elemento de su propio idioma- sino que existe una rectitud natural de las denominaciones, la misma para todos, tanto griegos como bárbaros. Yo le pregunto entonces si ‘Crátilo’ es en verdad su nombre o no; él está de acuerdo que sí. “¿Qué nombre le corresponde a Sócrates” –dije yo. “Sócrates”- dijo él. “Por tanto, ¿no le corresponde también a cada uno de todos los demás hombres el nombre con el que llamamos a cada uno?” “A ti sí que no te corresponde el nombre de ‘Hermógenes’ -dijo él- ni aunque todos los hombres te llamen así. (383 a-b)

Del pasaje anteriormente citado, se pueden señalar algunas consecuencias importantes que se presentan en la discusión de manera particular, comprometidas con la tesis naturalista.

De manera esquemática, éstas serían:

- i) dada la relación entre nombres y cosas nombradas, la rectitud (*orthóthēs*) de cada nombre (*ónoma*) va acompañada de la idea de cómo a través de los nombres, decimos algo verdadero del mundo.
- ii) dicha rectitud de cada nombre para cada cosa, se sostiene bajo el supuesto de una correlación entre elementos primitivos de nombres a cosas (las letras y sílabas, imitarían la esencia de cada cosa).
- iii) si asumimos que independientemente del idioma que se trate los nombres “significan” (pintan, imitan) cosas, por ende, la universalidad de los significados se sostendría si sabemos que tratan de una misma referencia (la realidad, tal cual es).
- iv) en vista que la rectitud de los nombres no ocurre por el acuerdo al interior de una comunidad lingüística particular se ha esquivado las consecuencias del relativismo cultural a saber: cada cultura a través de su propia lengua, constituye arbitrariamente la relación de nombres y cosas según los acuerdos a los que lleguen. Así, la verdad de lo que se diga sobre el mundo estaría fijada por el propio acuerdo, en cada cultura.

- v) con el ejemplo, de la “incorrección” del nombre, del propio Hermógenes¹, se deja claro que, aunque una comunidad lingüística emplee una denominación para hablar de una cosa, esto no significa, que expresen el ser de esa cosa, verdaderamente. El significado de los nombres, no dependerá de su uso.

Como dije al inicio de este artículo, la relación entre el lenguaje y la realidad que me interesa es aquella en la cual se tenga la capacidad de saber, que al usar el lenguaje se pueden decir cosas verdaderas del mundo. En la (i) afirmación, se supone que la rectitud de los nombres produce tal *lógos* verdadero. En la (ii) afirmación, se sugiere aquella condición de la rectitud de los nombres, a través de una correlación isomórfica entre primitivos y estructuras. La (iii) afirmación, defendería la objetividad de los significados, más allá de cualquier idioma, asumiendo la tesis ontológica del realismo que pondera la existencia de un mundo independiente y que es cognoscible objetivamente. Las afirmaciones (iv) y (v), son consecuencias de las tres primeras y discuten la veracidad de la tesis convencionalista de la rectitud por el acuerdo y los significados dados por el uso defendida por Hermógenes. Así mismo, expone su punto de vista:

Hermógenes: Yo por mi parte, Sócrates, ya he conversado sobre esto frecuentemente con él [con Crátilo] y con muchos otros, y no se me puede convencer de que haya alguna otra rectitud de la denominación que no provenga de la convención y del acuerdo. Pues a mí me parece que la denominación que alguien le pone a algo, ésa es la correcta; y si alguien a su vez la cambia por otra y ya no usa aquélla, la posterior no es menos correcta que la anterior; así como nosotros les cambiamos el nombre a los sirvientes, y el nombre cambiado no es menos correcto que el anteriormente dado. Pues por naturaleza no se ha producido ninguna denominación para cosa alguna, sino por convenio y por costumbre de quienes han creado esa costumbre y utilizan esa denominación. (384 d)

La tesis original de Hermógenes, no solo plantea la rectitud de los nombres según el acuerdo (*homología*), la convención (*synthékē*), el convenio (*nómos*) y la costumbre (*éthos*), sino que abraza, un relativismo

¹ El nombre ‘Hermógenes’, proviene de Hermes, dios del comercio y la ganancia, sin embargo, Hermógenes era pobre.

radical, en el que las cosas, son para cada quien, tal y como le parece que son, sin ningún punto estable que, como advierte Sócrates, tiene una clarísima deuda con el *homo mensura*, de Protágoras (Ibíd. 385 e).

Esta consecuencia ya fue señalada en la afirmación (iv), aunque referida al uso colectivo de los nombres al interior de una comunidad lingüística. Pero la posición original de Hermógenes oscila entre estos dos alcances, particular y colectivo de relativismo. Presionado por la clásica estrategia con la que Sócrates habitualmente discutía con los sofistas, es decir, aquella que les obligaba a establecer diferencias entre afirmaciones verdaderas y falsas (Ibíd. 385 b), Hermógenes, expone este doble relativismo:

Hermógenes: Pues yo al menos, Sócrates, no conozco otra rectitud de la denominación que ésta: yo puedo llamar cada cosa con una denominación que yo he dado, y tú con otra que tú por tu parte has dado. Así veo también que cada una de las ciudades tiene puestas denominaciones de forma particular a veces para las mismas cosas, lo cual ocurre a los griegos en relación con los demás griegos, y a los griegos en relación con los bárbaros. (Ibíd. 385 e)

Revisemos esto con más detalle. Para entender este asunto, llamemos *relativismo 1*, al tipo de relativismo particular en el que cada uno pone los nombres a su antojo y *relativismo 2*, al tipo de relativismo en el que son las comunidades, cada una por separado las que determinan la corrección de los nombres y sus significados, según el acuerdo o consenso, al interior de cada comunidad.

La primera modalidad de relativismo particular, sugeriría que si cada quien pone los nombres según su propio parecer, la corrección, sería relativa a cada persona en tanto que individuo. Sin embargo, esta posición relativista podría radicalizarse hasta volverla absurda, sin recurrir aún a la distinción de verdadero y falso. Hagamos de cuenta que en efecto, usamos una denominación distinta cada vez que nos referimos a las cosas mediante nombres. Como es del todo obvio, el solo hecho de emplear nombres para dar cuenta de las cosas, presupone que este empleo, tiene el propósito de comunicarse con algún otro interlocutor y no pretendería ser un mero acto adánico de poner nombres a las cosas por primera vez. Si esto es así, la dificultad salta a la vista. Tanto yo, como mi interlocutor, usamos nombres, pero siguiendo la tesis de

Hermógenes, los cambiamos continuamente y ninguno dejaría de ser correcto.

Independientemente de la corrección de tales denominaciones, si las mías o las suyas las serían, ocurriría que cuando yo mencione un nombre para referirme a una cosa, mi interlocutor, podría no entender a qué diantres me refiero, salvo que señale el objeto con todas las limitaciones que esto implicaría (asumir de plano que mi interlocutor interpreta que me refiero, en efecto, a un objeto completamente y no solo a una parte de éste o que el objeto al que me refiero, se ubica en un espacio determinado o también que simplemente, estoy hablando de objetos y no de alguna situación o hecho difícilmente localizable).

Considerando que no me da la gana de señalar el objeto nombrado en ese supuesto intercambio con mi interlocutor, tanto lo que él o ella me diga y lo que yo le diga, solo restringiendo el lenguaje al uso de los nombres para referirse a las cosas, podría no significar nada. Si tanto yo, como mi interlocutor, usamos nombres diferentes en todas las nuevas ocasiones, en las cuales nos refiramos incluso a los mismos objetos, simplemente no podríamos comunicarnos. En vista que soy yo, quien pone los nombres a las cosas, a mi parecer, inventando los nombres por mí mismo y de la misma forma sucede con mi interlocutor, por tal motivo, uno y otro, estaríamos continuamente, usando nuevas denominaciones que desconoceríamos, salvo las propias. Ocurriría, por ello, que bajo esta lógica en el *relativismo 1*, no solo, nada tendría un nombre correcto, sino que, prácticamente, no podríamos entendernos.

La segunda forma de relativismo colectivo, sugiere más bien, que compartimos un mínimo con el cual dentro de la comunidad a la que pertenecemos, denominamos a las cosas. Esta segunda forma de relativismo, anula inmediatamente a la primera, debido a que no sería cierto que cada quien, ponga los nombres según su propio criterio y cada vez que lo desee, porque para ser una comunidad, al menos, debe existir, un conjunto de nombres constantes, que conformarían nuestro uso habitual, el mismo que podría ser enseñado y aprendido. No obstante, como indicamos en la afirmación (iv) de nuestro esquema inicial, este *relativismo 2*, nos lleva a la consecuencia, que nuestro acuerdo sobre los nombres respecto de las cosas incluye también, supuestamente, la verdad de la corrección fruto del mero acuerdo entre las partes.

La observación principal es que, aparentemente, dicho acuerdo al interior de cada comunidad a su vez, podría modificarse, no pocas veces,

salvando ciertos mínimos para la comunicación, pero esto significaría, que la corrección de las denominaciones, junto con la verdad que acompaña a las mismas se estaría cambiando con cada nuevo ajuste en los acuerdos futuros.

En resumen, lo problemático aquí es que en el *relativismo 2*, la verdad dependería siempre del mismo acuerdo en cada comunidad (relativismo cultural²).

Formas y figuras: del *Crátilo* al *Tractatus*

Tras la crítica y refutación del relativismo manifiesto en la tesis inicial de Hermógenes, Sócrates intentará esta vez, convencer a su interlocutor que existe un modo en el que las cosas son según una esencia estable y que existen por sí mismas, tal como son por naturaleza (386 e). En adelante, Sócrates asumirá una defensa de la tesis naturalista de las denominaciones, hasta agotar sus posibilidades. Podría reconstruirse el argumento naturalista de Sócrates así (386 e-390 e):

- a) Si no es verdad, que las cosas son y existen para cada uno de modo particular (o para todos de manera igual, al mismo tiempo y siempre), entonces las cosas mismas, tienen una esencia (*ousía*) estable.
- b) Las acciones, son también una especie de lo que es, por tanto, las acciones, también se realizan, según su propia naturaleza.
- c) Hablar (*légein*), es una acción y hablar correctamente presupone el uso del instrumento naturalmente apropiado para decir las cosas.

² Una salida a esta dificultad es la que sugiere Hilary Putnam. En principio, desde la posición que él denomina "realismo interno" o realismo con "r" minúscula, se puede establecer una diferencia relevante, entre el relativismo y la "relatividad conceptual". Dice Putnam: "relatividad conceptual suena como "relativismo", pero no tiene ninguna de las implicaciones del relativismo del tipo "no existe ninguna verdad que descubrir"... "verdadero" no es más que un nombre para aquello en lo que un grupo de gente puede estar de acuerdo" (Putnam 61-62). Y concluye: "nuestros conceptos pueden ser relativos a una cultura, pero de aquí no se sigue que la verdad o la falsedad de cualquier cosa que digamos usando esos conceptos sea simplemente "decidida" por la cultura". (Ibíd. 64). Desde luego, el planteamiento de Putnam, no deja de ser problemático, sobre todo, en el aparente rechazo, a la tesis de la existencia de un mundo con independencia de nuestras descripciones, como parte de su crítica, a lo que él llama "realismo metafísico". Una muy sugerente observación a la idea de la relatividad conceptual putnamiana, es la planteada por Daniel Kalpokas (2007), en "Realismo, relatividad conceptual e independencia ontológica" (141-166).

- d) Denominar (nombrar), es una parte del hablar, por tanto, denominar es una acción.
- e) En vista que denominar es una acción comprometida con el hablar correcto, se infiere que, el instrumento del denominar es la denominación (el nombre).
- f) La denominación es un instrumento que enseña a distinguir las cosas, según su esencia.
- g) Aquel que enseña, usará correctamente la denominación; la usará "bien" (*kalôs*).
- h) Los nombres que usa, el que enseña, los ha obtenido por ley (*nómos*).
- i) Por ello, el que enseña usando las denominaciones, emplea la obra del legislador (*nomotetes*), aquel que es un forjador de palabras, un hacedor de nombres.
- j) El legislador, cuando hace un nombre, mira la forma (*eîdos*).
- k) El legislador, sabe poner en los sonidos y en las sílabas, la denominación por naturaleza apropiada a cada cosa, mirando, lo que es el nombre mismo.
- l) El legislador, deberá siempre dar a cada cosa, la forma del nombre que le conviene (o corresponde).
- m) Aquel que va a utilizar la obra del legislador, sabrá juzgarla una vez fabricada. Éste es el dialéctico el que sabe preguntar y responder.
- n) La obra del legislador, la de poner nombres a las cosas, deberá contar con la supervisión del dialéctico, si va a establecer las denominaciones correctas.
- o) En conclusión, Crátilo dice la verdad: los nombres corresponden por naturaleza a las cosas. El artífice de nombres, es aquel que mira hacia el nombre que corresponde por naturaleza a cada cosa y que es capaz de poner su forma en las letras (*grámmata*) y en las sílabas (*syllabás*).

El argumento de Sócrates que acabamos de reconstruir, presenta por lo menos un par de aspectos relevantes para nuestra presente discusión: A) en efecto, existe una correlación natural de nombres a cosas, en principio, mediante la *forma* de los nombres; B) dicha forma, deberá expresarse en los elementos primitivos del nombre. Ciertamente, la correlación de la que hablamos entre nombres y cosas requiere de lo afirmado en (a), que las cosas, tienen una esencia propia. Esto significa que los nombres se ajustarían a las cosas desde sus formas a las esencias de éstas. Pero algo más se ha indicado en B); a saber que hay elementos primitivos, con los cuales se conectan los nombres con las cosas.

Aunque es muy difícil sostener que existe una especie de correspondencia, punto por punto, entre los nombres y las cosas desde sus primitivos, no cuesta mucho advertir que alguna modalidad de este tipo de isomorfismo, está presente. Cuando menos, sabemos de antemano, por la argumentación reconstruida de Sócrates que hay elementos que componen a los nombres. Pero hace falta dar cuenta de un aspecto del carácter imitativo de los nombres, en tanto son pinturas de las cosas. El rasgo imitativo de las denominaciones, pone en evidencia que los nombres, deben ser semejantes a las cosas, incluso desde sus primitivos.

Dice Sócrates:

Sócrates: (...) las denominaciones no llegarían a ser semejantes a ninguna cosa, si primero estas cosas de las que se componen las palabras no tuvieran cierta semejanza con aquellas cosas de las que las denominaciones son imitaciones. (434b)

Si tomamos en cuenta la conclusión del argumento naturalista de Sócrates en la que afirmaba, que la forma del nombre es puesta en sus elementos primitivos son éstos los que, por ende, siempre que el nombre imite a la cosa como una pintura se conectan con la esencia de la cosa. Para advertir la relación que quiero establecer entre el *Crátilo* y el *Tractatus* desde una teoría figurativa, vale la pena hacer una extensión del argumento naturalista de Sócrates a propósito de los nombres como pinturas de las cosas. Dicha extensión, que podría llamarse *argumento pictórico*, sería la siguiente (Ibíd. 423 b-430 e):

- p) La denominación, es una imitación de aquello que imita.
- q) Las letras y las sílabas, en tanto elementos de los nombres, imitan la esencia de las cosas.
- r) Parece ridículo, que las cosas lleguen a ser claras, por haber sido imitadas por letras y sílabas, pero no tenemos un mejor recurso que éste.
- s) Al parecer el legislador en cuanto a las demás palabras, las refiere a letras y sílabas, creando para cada cosa, un signo y un nombre, a partir de los que, imitando, compuso también el resto mediante estas mismas sílabas y letras.
- t) La imitación de los nombres, no es solo correcta, sino también, verdadera, cuando se refiere a lo semejante.
- u) La denominación es una imitación como una pintura.

- v) En conclusión: los nombres llegan a ser semejantes a las cosas, en tanto son pinturas, cuando sus elementos primitivos, imitan la esencia de las cosas.

Hagamos un apretado resumen de lo afirmado hasta aquí: desde la tesis naturalista defendida por Sócrates los nombres son correctos, toda vez que imiten a modo de pinturas (o figuras) la *esencia* de las cosas, pero esta imitación se produce desde sus elementos primitivos en los que está puesta la *forma* de los nombres. A su vez, las letras y sílabas constituyen la base desde las cuales se forja los demás nombres. Esta es la teoría figurativa del *Crátilo*, sobre la que me referiré a continuación en el *Tractatus*.

Habría que señalar, desde un inicio, que algunas de las analogías que aquí se plantean, son solo parte de una lectura que espero al menos pueda resultar creativa. Si resulta que alguna de sus partes, parece forzada, lo más probable, es que así sea. Después de todo, de lo que se trata es de cómo leer un texto filosófico, en este caso, un clásico comprometido con problemas contemporáneos de modo que pueda a su vez, ayudarnos a abrir horizontes de problemas novedosos y respuestas imaginativas. Veamos algunos puntos de encuentro y luego sus desajustes.

Una de las formas en la que puede leerse el *Tractatus* y por medio de la cual relacionamos al lenguaje con el mundo, corresponde a lo que se conoce como la forma general de la proposición que atraviesa la propia teoría figurativa de la realidad desde el plano de la proposición con sentido. Como es del todo imposible entrar en detalles, procuraré señalar dichas “coincidencias” con el diálogo platónico. De manera semejante a como se hizo con los argumentos de Sócrates, haré una reconstrucción de las afirmaciones del *Tractatus*, que refieren a la figura y la forma de la figuración del mundo desde el discurso proposicional:

- 1) El mundo (*Welt*), está compuesto de hechos (*Tatsachen*).
- 2) El estado de cosas (*Sachverhalten*), es una conexión de objetos (cosas).
- 3) El objeto es simple.
- 4) En los estados de cosas, los objetos, están unidos entre sí, como los eslabones de una cadena.
- 5) Hacemos figuras de los hechos.
- 6) La figura (*Bild*), representa el estado de cosas en el espacio lógico.
- 7) Los elementos de la figura, hacen en ella, las veces de los objetos.
- 8) La interrelación de los elementos de la figura se llama su estructura

- y la posibilidad de la misma, su forma de figuración (*Form der Abbildung*).
- 9) La relación figurativa, consiste en las coordinaciones entre los elementos de la figura y los elementos de las cosas.
 - 10) Lo que la figura, ha de tener en común con la realidad para poder figurarla a su modo y manera -correcta o falsamente- es su forma de figuración.
 - 11) Si la forma de la figuración, es la forma lógica (*Logische Form*), la figura se llama, la figura lógica (*Logische Bild*).
 - 12) La figura, tiene en común con lo figurado, la forma lógica de la figuración.
 - 13) La proposición es una figura de la realidad.
 - 14) Los signos simples usados en la proposición se llaman nombres.
 - 15) El nombre significa el objeto.
 - 16) Solo la proposición tiene sentido; solo en la trama de la proposición, tiene un nombre significado (*Bedeutung*).

El discurso proposicional, configura un mundo con sentido, siempre que los hechos que conforman el mundo se correlacionen con las proposiciones (o los estados de cosas, con las proposiciones elementales), a través de una forma común, una forma lógica o forma de la figuración. Los nombres, como elementos primitivos de las proposiciones se conectan con los primitivos en el plano del mundo, esto es, con las cosas, siendo éstas sus significados. Por tanto, se halla un modelo atomista, en el plano figurativo del mundo en que se conectan, tanto estructuras, como elementos.

El plano del sentido, es el plano proposicional y con esto, habría una relativa prioridad de las proposiciones sobre los nombres a nivel cognitivo, porque más tarde, es el sentido de las proposiciones comparadas con la realidad, lo que permite un discurso verdadero. Esto es, la verdad o la falsedad se predicen de las proposiciones, no de los nombres en tanto primitivos.

Sin embargo, es verdad que la proposición figura un hecho, lo representa, lo describe. No voy a considerar aquí el aspecto constructivo, activo de *Bild*. Para señalar los puentes con el *Crátilo*, es suficiente saber que así como los nombres “pintan” a las cosas, las proposiciones “figuran” y así, como los nombres tienen primitivos, las proposiciones, también. Así mismo, como hay una forma de los nombres que puesta en los elementos, pretende imitar la esencia de las cosas, también existe una

forma de la figuración, que compromete al lenguaje y al mundo en términos representacionales.

Hay un esquema atomista e isomórfico en ambos contextos, aunque no tienen el mismo alcance, pero uno y otro esquema contribuyen a un mismo fin: figurar, pintar, representar, ya sea con nombres o con proposiciones, lo que significa que en uno y en otro contexto, el lenguaje es asumido principalmente, para este fin representacional.

Vemos así, que bajo un mismo modelo figurativo representacional, el *Crátilo* y el *Tractatus*, asumirían la primera respuesta a cómo se relaciona el lenguaje y el mundo; esto es, es posible conocer el mundo desde el lenguaje, gracias a una conexión especial compartida. Aunque claro, ésta no es toda la historia.

Reflexión final: más allá del lenguaje

Cuando Sócrates discute con Crátilo, hacia el final del diálogo, sobre las propias consecuencias de la teoría naturalista de las denominaciones, observa que, probablemente, no han estado siguiendo el método correcto, porque pareciera, dice Sócrates, que empezar por los nombres y no por las cosas conduce a un error. Esto significaría, que no se conoce la esencia de las cosas desde los nombres, sino desde las cosas mismas, razón por la cual si estuviera en nuestras manos iríamos “fuera de las palabras”, para conocer las cosas sin palabras (438 d-e).

El recelo planteado por Sócrates, respecto de cómo el lenguaje podría sortearse con el ánimo de conocer verdaderamente las cosas del mundo, sumado al hecho de que el lenguaje, aparece más como un disfraz de la realidad y un medio que produce el engaño más que la verdad es como dije al inicio de este artículo, una de las formas en que también podía contestarse a la pregunta sobre la relación del lenguaje y el mundo.

Se trataría de ir más allá del lenguaje. Como una última conexión entre el universo del *Crátilo* y el *Tractatus*, podría sugerir, cómo es que Wittgenstein, cuando reflexiona sobre el ámbito místico, comprometido con el sentido del mundo y la ética (6.41 6.44 6.521 6.522), sugiere,

la apuesta por ir más allá del lenguaje³, del lenguaje figurativo representacional, debido a que el sentido, el *qué es*, tendría que estar fuera del mundo; y como los límites del mundo, son los límites del lenguaje, esto no sugeriría sino que, para los problemas vitales, éstos que no son siquiera tocados luego de agotar la figuración del mundo (6.52), habría que ver y no decir (6.521), al menos, no decir en los términos figurativos en que el mundo se reduce a un *cómo es*, de tipo causal explicativo. Cómo es que podríamos comprender este repliegue silencioso que da cuenta del sentido del mundo, es un misterio; cómo es que podríamos compartir nuestra visión del sentido del mundo, suponiendo que se haya alcanzado es un misterio, aún más profundo.

REFERENCIAS BIBLIOGRÁFICAS

Kalpokas, Daniel. "Realismo, relatividad conceptual e independencia ontológica". *Analítica*. 2007: 141-166. Impreso.

Platón. *Crátilo o del lenguaje*. Madrid: Editorial Trotta, 2002. Impreso.

---. *Crátilo*. México: UNAM, Bibliotheca Scriptorum Graecorum et Romanorum Mexicana, 2008. Impreso.

Putnam, Hilary. *Las mil caras del realismo*. Barcelona: Editorial Paidós, 1994. Impreso.

Wittgenstein, Ludwig. *Tractatus Logico Philosophicus*. Madrid: Alianza Editorial, 2001. Impreso.

---. "Conferencia de ética". *Obras Wittgenstein II. Diarios y conferencias*. Trad. Fina Birulés. Madrid: Editorial Gredos, 2009. Impreso.

³ La misma idea la encontramos en la "Conferencia sobre ética": "veo ahora que estas expresiones carentes de sentido no carecían de sentido por no haber hallado aún las expresiones correctas, sino que era su falta de sentido lo que constituía su mismísima esencia. Porque lo único que yo pretendía con ellas era, precisamente, *ir más allá* del mundo, lo cual es lo mismo que ir más allá del lenguaje significativo. Mi único propósito -y creo que el de todos aquellos que han tratado alguna vez de escribir o hablar de ética o religión- es arremeter contra los límites del lenguaje" (Wittgenstein 523).

COMPRENSIÓN Y LENGUAJE*

UNDERSTANDING AND LANGUAGE

MAURICIO VÉLEZ UPEGUI

Universidad Eafit, Colombia. mavelez@eafit.edu.co

RECIBIDO EL 28 DE MARZO DE 2012 Y APROBADO EL 18 DE MAYO DE 2012

RESUMEN ABSTRACT

El presente artículo pretende exponer, respecto de las posturas hermenéuticas de Gadamer y Ricoeur, el modo como ambos pensadores parecen responder al problema que interroga por el papel que desempeña el lenguaje en la actividad humana de la comprensión, ya en el ámbito del diálogo, ya en el de la relación que media entre el acto de leer y el acto de escribir.

This paper intends to show, regarding the hermeneutic views of Gadamer and Ricoeur, the way how both thinkers seem to respond to the problem that inquires for the role that language plays in the human activity of understanding, both in the field of dialogue, and in the relation that mediates between the act of reading and the act of writing.

PALABRAS CLAVE KEY WORDS

Comprensión, diálogo, escritura, modelo de preguntas y respuestas, lectura, lenguaje.

Understanding, dialogue, questions and model answers, language, reading, writing.

* Este artículo es el resultado de la investigación: "Comprensión y lenguaje", adscrita al grupo de investigación "Estudios sobre política y lenguaje" del Departamento de Humanidades, de la Universidad EAFIT, desarrollada entre enero y diciembre de 2010.

Comprensión y lenguaje

El propósito del presente artículo es presentar, mediante una exposición condensada, algunos de los resultados de la investigación titulada "Comprensión y lenguaje", cuyo objeto de estudio pretendía comparar y contrastar la hermenéutica fenomenológica de Gadamer y la fenomenología hermenéutica de Ricoeur (Grondin, *El legado* 114ss).

Entre otros objetivos trazados, la investigación en mención se proponía dar respuesta a una pregunta puntual, a saber: ¿qué papel desempeña el lenguaje en la actividad humana de la comprensión, ya en el ámbito del diálogo, ya en el de la relación que media entre el acto de leer y el de escribir? Fundábamos dicha pregunta en la conjetura de que tanto el pensador alemán como el francés hacían a eco a la idea heideggeriana de que los seres humanos somos en el mundo, estamos en él, vivimos mundanamente, de una manera comprensiva. Y, de partida, lo hacemos, comprendemos, nos disponemos a comprender, no por antojo o capricho, sino porque la comprensión es una posibilidad inherente a la existencia humana misma (Heidegger, *Ontología* 32ss).

En las líneas que siguen, por tanto, procederemos formalmente de la siguiente manera: en el numeral 1, el diálogo y no la relación lectura-escritura, será el tema de exposición, conforme a un orden que hace aparecer primero a Gadamer y luego a Ricoeur; en el numeral 2, al contrario, la relación lectura-escritura, y no el diálogo, será el asunto de reflexión, según un orden que hace intervenir primero a Ricoeur y después a Gadamer. Sirviéndonos de un símil propio del quehacer pictórico, diríamos que hemos articulado el texto como si se tratara de un cuadro regulado por la antisimetría. Por último, en el numeral 3, cerraremos lo dicho, pero no sin aludir a la advocación heideggeriana que nos ha servido de punto de partida.

No sobra acotar que, aparte de los textos básicos de los autores indicados, nos serviremos de diversas fuentes secundarias, las cuales han contribuido con sus voces explicativas, si no a resolver, por lo menos a matizar convenientemente el problema.

I

Tomemos en primer lugar el discurso oral y, sobre todo, la realidad verbal donde éste se materializa, a saber: el diálogo.

Tanto para Gadamer, como para Ricoeur, el diálogo constituye la quintaesencia de la comunicación humana. Diálogo, no sólo en el sentido de un encuentro conversacional entre dos individuos que se disponen a “hablar unos con otros” en la esperanza de alcanzar una comprensión común del mundo, sino también en el sentido de un encuentro -a menudo inconsciente- entre nosotros mismos y la tradición de la cual participamos. Si el primero concibe el diálogo como evento esencialmente humano en el que el lenguaje se realiza vivamente (Gadamer, *Verdad II* 203), el segundo afirma que en él los seres humanos encuentran una condición existencial para intentar “transgredir o superar su soledad fundamental”, es decir, aquella soledad derivada de la imposibilidad de transferir directamente lo experimentado a alguien más (Heidegger, *Ontología* 30).

Siguiendo a Gadamer, anotemos que la lógica que gobierna la experiencia dialógica, no es la lógica “enunciativa” que rige la formalización matemática de la ciencia. En ésta no cuenta más que lo enunciado (el *logos apophantikós*), cuyo

único sentido es realizar el *apophainesthai*, la auto-manifestación de lo dicho. Es una proposición teórica en el sentido de que abstrae de todo lo que no se dice expresamente. Sólo aquello que ella misma manifiesta con su enunciación constituye aquí el objeto de análisis y el fundamento de la conclusión lógica. (Gadamer, *Verdad II* 189)

No obstante, en la vida diaria no tendemos a emitir esta clase de enunciados. Muy al contrario, nuestra forma de hablar, nuestra forma de hacer frente al pasado, tiende a ser esquemática y con frecuencia, plagada de equívocos e imprecisiones que, ¡oh misterio!, no causan, sin embargo, una total incomunicación. Y aún si la causaran, contamos con la posibilidad, tramitada por nuestra conciencia lingüística y más, por nuestra conciencia comunicativa (que incluye la dimensión no verbal, una de las formas de entender la noción de “lingüisticidad”), de apelar al recurso de las preguntas y respuestas. De ahí que para Gadamer sea la

“lógica” de las preguntas y respuestas la que domine nuestra inserción en el diálogo que somos o, lo que es igual, nuestra inserción en el mundo.

Además, no hay diálogo, según Gadamer, que no implique la función del habla. Y es ésta la que lo lleva a definirla “como un hacer patente lo no actual por medio del lenguaje, de modo tal que otros lo puedan ver” (Ibíd. 145) y a postular la idea -que haría parte de una ontología del lenguaje- según la cual

el que habla un idioma que ningún otro comprende, en realidad no habla... En ese sentido el habla no pertenece a la esfera del yo, sino a la esfera del nosotros. (Ibíd. 150)

No cabe duda de que tanto la definición de habla propuesta por Gadamer como su postulado ontológico guardan relación con el carácter existencialmente discursivo de los seres humanos. Si hablar equivale, según lo expuesto antes, a hacer discursos y si el acto de hacer discursos -bajo la determinación de unidades predicativas que se materializan en oraciones o enunciados- tiene por finalidad la comunicación dirigida a otros que coexisten originariamente conmigo en el mundo, entonces no hay habla que no se abra discursivamente al otro para hacer ver, para hacer comprensible, no tanto lo que “causa placer, para buscarlo, y lo que causa dolor o daño, para evitarlo” (debido a que este sería sólo el caso de los animales), sino más bien para hacer ver lo que es útil o beneficioso y lo que es injusto o dañino. Y más todavía, para

pensar lo común, tener conceptos comunes, sobre todo aquellos conceptos que posibilitan la convivencia de los hombres sin asesinatos ni homicidios, en forma de vida social, de una constitución política, de una vida económica articulada en la división del trabajo. (Gadamer, *Verdad II* 145)

No obstante, por lo que atañe al ser del discurso, ¿acaso al hablar tenemos conciencia del modo -lexical, sintáctico y semántico- como lo hacemos? Ciertamente que ante determinadas situaciones de producción discursiva en las que el otro coexiste físicamente con nosotros y ante quien decimos algo sobre algo, hacemos consciente el discurso que pronunciamos. De hecho, si bien es cierto que a cualquiera podemos hablarle, es siempre el otro quien motiva el despliegue de nuestra cadena discursiva. Dado que el otro comparece ante nosotros como portando una determinada investidura, un determinado rol social, obramos discursivamente de

conformidad con dicha investidura, con dicho rol. Por ello, tomamos todas las precauciones lingüísticas del caso. Decidimos emplear, por ejemplo, en relación con las fórmulas de tratamiento, un pronombre personal en vez de otro, en el entendido de que los pronombres comportan una semántica de la familiaridad o del distanciamiento; si vacilamos, en lo atinente al empleo correcto de un término, preguntamos al otro “¿es adecuada esta palabra?”, espoleados por la esperanza de que su autorización nos mueva a continuar hablando, entre otros. Así las cosas, el fantasma del otro regula por lo general nuestros comportamientos lingüísticos.

Pero, no es menos cierto, a juicio de Gadamer, que lo propio del lenguaje, lo característico del ser del discurso, es “su auto-olvido esencial” (Gadamer, *Verdad II* 149). Auto-olvido significa esto: mientras hablamos, como sujetos de discurso, arrastrados por el orden estructural que impone el sistema de la lengua de la cual somos tributarios, reparamos, excepcionalmente, en la manera como articulamos el entramado significativo de nuestro discurso. Dicho con otros términos: mientras ejecutamos la acción verbal de predicación discursiva, no pensamos en si estamos o no repitiendo una palabra que acabamos de pronunciar, o en si el tipo de oración que emitimos es yuxtapuesta o subordinada adverbial de tiempo o en si el tiempo verbal que rige la entrada de una frase condicional, debe ser formulada en modo indicativo o subjuntivo. Y no pensamos en nada de eso, no sólo porque al hablar nos concentramos más en el *sobre-qué* del discurso que en su configuración formal, sino porque, además, lo más probable es que no sepamos nombrar técnicamente las distintas unidades léxicas, sintácticas o semánticas que sirven de vehículo a nuestra expresión. Una inconsciencia general del lenguaje, entonces, envuelve cada una de nuestras actuaciones verbales.

De querer hacer consciente la naturaleza, estructura y funcionamiento del lenguaje que hablamos, haría falta, puntualiza Gadamer, una enorme labor abstractiva, porque “el lenguaje real y efectivo desaparece detrás de lo que dice en él” (Gadamer, *Verdad II* 149). Sabemos que hablamos, pero, a menudo, no sabemos cómo lo hacemos, ni mucho menos cuáles son las reglas que seguimos al momento de hacerlo. Asimismo, nada garantizaría que de llegar a conocer tales reglas modificaríamos sustancialmente nuestra particular manera de comportarnos discursivamente. De cierta manera, el fenómeno que acabamos de describir explica el carácter mágico del lenguaje.

Si no fuera porque, por alguna razón misteriosa, no llegamos a transgredir completamente el complejo de reglas que subyacen a nuestras realizaciones lingüísticas, diríamos que no habría razón alguna para invocar el “auto-olvido” esencial del lenguaje como argumento a tener en cuenta a la hora de vislumbrar algunas relaciones entre la comprensión y el lenguaje. Pero es justamente porque la desatención en el *cómo* de la realización lingüística puede afectar seriamente la configuración de sentido del *sobre-qué* de la misma, por lo que mencionamos dicho aspecto como un elemento que es preciso tener presente a la hora de considerar dicha relación.

Desde la perspectiva hermenéutico-filosófica agenciada por Gadamer, es posible afirmar que comprender es “captar lo que nos tiene prendidos” (Grondin, *Introducción* 39). Si digo a alguien “comprendo”, digo no sólo que “agarro” algo sino, también, cosa menos consciente, que *soy agarrado por la cosa*. Nadie que afirme que comprende algo y nadie de quien se predique que es una persona entendida en algo, puede alegar que como sujeto permanece separado de su objeto.

Desde esta perspectiva, comprender es un proceso que implica un doble movimiento: comprendo siempre algo y ese algo me *prende*, me *trama*, me *arropa*. Lo curioso es que aunque a menudo decimos que comprendemos, a menudo también somos incapaces de establecer la manera como comprendemos y a veces, aunque parezca paradójico, el contenido mismo de lo comprendido. Razón le asiste a Gadamer cuando asevera que el comprender es un enigma; y más, “un enigma del cual vivimos”. Prueba de ello es que nosotros mismos o alguien más podría acosarnos con esta pregunta incómoda: “¿Qué es lo que propiamente se entiende, cuando se entiende?” (Ibíd.).

Ya lo planteaba Bauman: si la comprensión es importante para el ser humano, no es porque ella sea relevante por sí misma, sino porque la incompreensión se manifiesta allí donde quiera que los seres humanos entran en contacto, se trenzan en conversaciones sucesivas o se consagran al arduo oficio de la escritura y la lectura. En efecto, ya en forma de lamento cotidiano -“No te entiendo”-, ya en forma de interpelación interrogativa -“¿Qué me quisiste decir?” (Bauman 122)-, bien en forma de estupefacción coloquial -“No entendí ni jota”-, bien en forma de perentoria admonición -“¡No confunda las cosas! ¡Eso no es lo que dice el texto!”-, la incompreensión, en cualquiera de sus diversas manifestaciones, llámense malentendidos, comprensiones a medias,

hermetismos, ausencia de entendimiento, entre otros, constituye un problema esencialmente lingüístico.

Las causas que la producen pueden ser múltiples y obran por igual cuando se trata de intercambios conversacionales o de aproximaciones de lectura a textos escritos. Respecto del primer ámbito, las causas, entre otras, pueden ser las siguientes: la distancia social entre los interlocutores; los juegos verbales en que puede incurrir alguno de los participantes; el uso de una sintaxis irregular por parte de alguno de los ejecutantes; la celeridad con que alguno de los querellantes encadena la lógica de sus enunciados y destila sus juicios; el énfasis puesto -por parte de quien escucha-, no en el enunciado total proferido, sino en unas cuantas palabras que hacen parte del enunciado, entre otros; y, en relación con el segundo, éstas y otras pueden ser las causas: el distanciamiento temporal que media entre el texto leído y la situación contemporánea del lector; la carencia de enciclopedia en el lector; la obstinación por parte del lector en tratar de imponerle al texto, como anticipaciones de sentido, sus propias ocurrencias; la dimensión pluri-significativa del lenguaje, etc.

Éstas y otras causas que dejamos de mencionar, acaso tengan todas algo en común. ¿Qué? Sirviéndonos de una expresión cara a Rafael Echeverría, diríamos que es la inexistencia de un “dominio consensual”¹ o la perturbación del “dominio consensual” existente en un momento determinado.

Si comprendemos y somos comprendidos por otros es porque nos movemos mundanamente en un horizonte de comprensión donde operan un sinnúmero de objetos de acuerdo implícitos regulados por una discursividad mundana. Diríase que la comprensión, como manifestación del *cómo hermenéutico* o existencial, aparece construida sobre la base de lo *ya-interpretado* que hace posible la realización de

¹ “Es en la interacción entre diferentes seres humanos particulares -antes incluso de que podamos hablar de un proceso de individualización en el que nos constituimos como personas- donde aparece una precondición fundamental del lenguaje: la constitución de un *dominio consensual*. Hablamos de consensualidad dondequiera que los participantes de una interacción social comparten el mismo sistema de signos (gestos, sonidos, etc.) para designar objetos, acciones o acontecimientos en orden a coordinar sus acciones comunes. Sin un dominio consensual no hay lenguaje... El dominio consensual se constituye en la interacción con otros en un espacio social” (Echeverría 50).

ulteriores actos de convivencia humana². Esa discursividad mundana, compuesta por los discursos que al tiempo que hablan sobre el mundo lo *constituyen*, es la que los seres humanos “aprenden” al hablar. Si nadie en particular “aprende” a hablar su lengua materna, ya que de hecho lo que ocurre es que la habla, todos aprendemos a usar lo consensual general contenido en la discursividad mundana. ¿Cómo se produce dicho aprendizaje de lo consensual general?

Para responder, Gadamer apela a la imagen aristotélica del ejército en fuga:

¿Cómo llega a detenerse un ejército que está en fuga?
¿Dónde ocurre que empiece a detenerse? Desde luego, no por detenerse el primer soldado, o el segundo, o el tercero. No se puede afirmar que el ejército se haya detenido por haber dejado de huir un determinado número de soldados en fuga. Porque con él no empieza el ejército a detenerse, sino que empezó a hacerlo desde mucho antes. Cómo se inicia el proceso, cómo se propaga, cómo finalmente, en algún momento, el ejército se detiene, es decir: obedece de nuevo a la unidad de mando, todo ello nadie lo dispone a conciencia, nadie lo controla según plan, nadie lo certifica cognoscitivamente. Y, sin embargo, indudablemente ha ocurrido. Algo análogo ocurre con el conocimiento de lo general. (Gadamer, *Verdad II* 148)

Si es dicha apropiación de lo consensual general la que nos proporciona un horizonte de comprensión, la alteración de lo consensual general o su inexistencia ayudaría a relevar el fenómeno de la incomprensión.

² El comprender que signa nuestra existencia no es un comprender vacío de interpretación, sino un comprender que arrastra consigo el carácter de *ya-interpretado*. Mal haría el hombre en acoger la creencia de que al ser arrojado al mundo, al encontrarse situado en él, el mundo brota y se manifiesta en una especie de esplendor virginal, ajeno a cualquier forma de captación preliminar. Para Heidegger, al revés, es inherente al ser de la facticidad el contar con interpretaciones precedentes o históricamente anteriores que forman parte de la vida misma. A resultas de este *ya-interpretado*, *pre-interpretado* o *comprender anticipado*, las cosas del mundo y los otros, aunque se nos resistan en la mostración de su ser más esencial, no se nos revelan fenomenológicamente, sin embargo, como completamente extrañas. Dado que al ser arrojados quedamos situados en el mundo, con otros y entre otros, con otras cosas y en medio de las cosas, comprendemos, sin necesidad de explicaciones teóricas o epistemológicas ulteriores, si no la razón de ser de los otros y las cosas, el para qué están ahí, bien como útiles que “están a la mano” para ser usados por nosotros en nuestro cotidiano vivir, bien sea como entes intra-mundanos a disposición para el trabajo de la ciencia o la técnica (Heidegger, *Ser y Tiempo* 94).

A manera de ilustración, ¿estaríamos dispuestos a admitir que así como nos empeñamos en auto-comprendernos en términos de seres racionales y, por consiguiente, como presuntos soberanos del lenguaje, estamos en mora de comprender que también somos vasallos de él? ¿Por ventura el conocimiento teórico del funcionamiento sistémico del lenguaje, según las conquistas incontestables de la lingüística estructural, nos ponen a salvo de incurrir en malentendidos? Más todavía: ¿será que por más cuidado que ponga el hablante en aquello que habla, y sobre aquello que habla, es capaz de conjurar la experiencia hermenéutica que nos hace saber que es siempre el otro, con su escucha, el que asigna el sentido a lo dicho? (Von Foerster 61).

Resumiendo: en el marco dialógico del que se ocupa Gadamer, siempre contamos con la posibilidad de apelar al primado hermenéutico de la pregunta y la respuesta (Gadamer, *Verdad I* 87-7) para intentar disolver la opacidad que instala todo fenómeno de incompreensión. Al estar en co-presencia física, los interlocutores tienen a la mano uno de los caracteres del ser del discurso, a saber, su recursividad o, lo que es igual, la posibilidad de servirse del discurso para interrogar el *sobre qué* del discurso y lo *dicho como tal* por él.

Y ello por no hablar del hecho de que la misma co-presencia física, aunque no constituye una manifestación discursiva en propiedad, admite ser comprendida e interpretada como si se tratara de una comunicación discursiva. Las miradas, los gestos, los ademanes, el espacio donde se ubican los interlocutores en realidad de verdad no hablan, pero es *como si* lo hicieran. Y eso que “dicen” de modo no verbal, eso que “expresan” de un modo indirecto, pero igualmente eficaz, puede complementar lo dicho verbalmente y puede ayudar, porqué no, a clarificar el momento de la incompreensión, restituyendo de esa manera el consenso alterado o creándolo en el acto comunicativo.

Algo similar propone Ricoeur. Atrás anotábamos que para éste el carácter dialógico del discurso constituía la base de cualquier investigación existencial, dada la soledad radical del ser humano. ¿Qué entender por ello?

Por soledad no me refiero al hecho de que frecuentemente nos sentimos aislados en una multitud, o al de que vivimos y morimos solos, sino, en un sentido más radical, a que lo experimentado por una persona no puede ser

transferido íntegramente a alguien más. Mi experiencia no puede convertirse directamente en tu experiencia. Un acontecimiento perteneciente a un fluir del pensamiento no puede ser transferido como tal a otro fluir del pensamiento. Aun así, no obstante, algo masa de mí hacia ti... Este algo no es la experiencia tal como es experimentada, sino su significado. (Ricoeur, *Teoría* 30)

Lo comunicado no es la vivencia que da cuenta de la experiencia, sino el sentido mismo de los enunciados que constituyen el diálogo. Sólo que éstos, al acontecer en el tiempo, no escapan a la evidencia de una realidad incontestable: la polisemia originaria del lenguaje.

Según Ricoeur, fuera del uso lingüístico prácticamente no existe expresión que no sea depositaria de múltiples sentidos. El espectro semántico se reduce si y solo si el término es usado en una situación comunicativa concreta. En esa medida la situación se torna creadora de un contexto comunicativo. Merced a esta creación,

la polisemia exige como contrapartida el papel selectivo de los contextos para poder determinar el valor semántico que toman las palabras en un mensaje determinado (Ricoeur, *Del texto* 72)

La creación de contextos, en medio de todo evento dialógico, suscita entre los interlocutores una actividad de descodificación (de "desciframiento"), que se basa en un modelo especial: el de las preguntas y respuestas. Y esa actividad es justamente la de la comprensión o, si se quiere, la de la interpretación: "reconocer qué mensaje relativamente unívoco ha construido el hablante sobre la base de un léxico regido constitutivamente por la polisemia" (Ibíd. 73).

En el caso de la polisemia, el *qué* de lo dicho puede ser interrogado respecto de su *querer-decir* con el fin de volver unívoco aquello que no lo es; y en el caso de la referencia, el *sobre qué* de lo dicho puede ser interrogado a partir de su plena significación enunciativa con el objeto de que el acto de "remitir a la realidad equivalga a remitir a *esta* realidad" (Ibíd. 103). Y aún si la lógica de las preguntas y respuestas se revelara insuficiente para resolver los problemas de comprensión suscitados por la polisemia o la referencia, el lenguaje, al acontecer como discurso, ofrece, al extremo, un último recurso de desciframiento: la dimensión ostensiva o digital. Si los pronombres demostrativos, los adverbios

de lugar y las variaciones pronominales constituyen “expresiones esencialmente ocasionales” (Husserl 272ss), es porque en ellos la frontera que delimita el carácter indicativo y expresivo de los signos aparece siempre borrosa, pues al tiempo que señalan, casi como si se trataran de gestos no verbales, ellos indican en cada caso, existencialmente, al sujeto mismo de la enunciación. Razón le asiste a Ricoeur cuando afirma que en el diálogo “el sentido muere en la referencia y la referencia muere en la mostración” (Ricoeur, *Del texto* 129-30).

En fin, si vivir de un modo hermenéutico consiste en mantenerse bajo el primado de acuerdos tácitos, en restituirlos toda vez que tomamos conciencia de que ya no los tenemos a la mano o en crearlos cuando sea menester hacerlo, conviene anotar que existe una diferencia notable entre la situación comunicativa inherente al diálogo vivo y la que concierne a los textos.

II

En efecto, ¿qué sucede cuando la realización discursiva del lenguaje deja de ser dialógica y se torna escrita? ¿Acaso escribimos como hablamos? Si la escucha es al tiempo condición de posibilidad del discurso y destino comunicativo de éste, ¿la lectura sería condición de posibilidad de la escritura y destino de su coronación interpretativa? Aunque Ricoeur, al tenor de una hermenéutica expandida, parece concebir una noción de texto de una manera bastante amplia (ya que a lo largo de su trayectoria intelectual insistirá en pensar la acción humana en términos de texto, o incluso los acontecimientos históricos como manifestaciones textuales y aún la identidad de la propia vida a semejanza de un texto), queda claro en él que hay primariamente texto cuando y donde el discurso se realiza de modo escrito.

Texto, especificará Ricoeur categóricamente, “es discurso fijado por la escritura” (Ricoeur, *Del texto* 128). A ese decir escrito que deviene texto le son correlativas dos características que no atañen al diálogo. Primera: si en el diálogo el discurso está dirigido a otro, que funge de interlocutor presente, y más, de interlocutor *presentificado* discursivamente por el carácter convencional de los pronombres que regulan la situación, al punto de que es la dinámica de la situación misma la que favorece, no sólo el intercambio de los roles conversacionales, sino la actualización de la lógica de la pregunta y respuesta, en el texto el discurso está dirigido a Otro (así, con mayúscula), que vale, no como interlocutor localizado,

presente en situación y pronominalmente *presentificado*, sino como eventual lector, singular o plural, siempre anónimo y siempre diferido. Mientras que en el diálogo la tríada emisor-mensaje-interlocutor implica la co-presencia física de los agentes comunicativos, así como la materialización oral de los enunciados, en la escritura dicha co-presencia no es completa: quien escribe no tiene delante de sí al lector, ni éste, al leer, tiene delante de sí a aquél. El primero, durante el proceso de escritura, *ve* cómo su texto se va haciendo poco a poco, aunque sin saber qué clase de lector se ocupará de su labor; el segundo, a poco de empezar la lectura, ve el texto terminado, aunque sin saber todavía cuáles son las articulaciones internas sobre las cuales se asienta ni el sentido que propone el texto elaborado por el “autor”.

Segunda: “Cuando el texto toma el lugar del habla, ¿qué pasa con la función referencial del lenguaje?” (Ricoeur, *Del texto* 129-39). En los textos, a diferencia del diálogo, no hay quien (locutor-“autor”) pueda resolver un problema de atribución de referencia. En ellos la referencia no se suprime, pero sí se pone temporalmente entre paréntesis. ¿Por qué? Porque la suspensión temporal de la relación del discurso con el mundo, con la realidad, reclama una articulación previa, necesaria para que pueda emerger su textura, su calidad misma de texto.

En Ricoeur esa articulación previa tiene que ver con la noción de obra. El texto como obra es una noción que entraña tres elementos conexos: lingüístico, estructural e individual. Lingüístico, porque decir texto equivale a decir conjunto de enunciados cuyas relaciones de cohesión sintáctica y coherencia semántica constituyen una unidad superior a la oración. Estructural, porque el texto, al poder entrar en contacto con otros textos, no puede menos de responder a las leyes de un canon discursivo. E individual, porque el texto, al ser el producto de una exteriorización intencional por parte de un sujeto de enunciación, recibe las improntas singulares del individuo que lo agencia. Los tres elementos dan al concepto de composición textual un sentido complejo, no exento de vínculos con el hecho de que toda obra, así considerada, es menos el resultado de un raptó de inspiración creadora que de un solícito esfuerzo intelectual. Porque el texto deviene obra, no resulta inútil establecer tipologías textuales y marcos de comparación estilísticos. El lector, por extensión, queda implicado. Éste, pese a ser anónimo, se convertirá en un destinatario privilegiado para el autor a condición de que realice una tarea hermenéutica. ¿Cuál? Dicho en palabras de Ricoeur (que en esto hace eco a Gadamer), la tarea consistente en establecer la “proposición de

mundo” que el texto formula a su manera. El esfuerzo de interpretación del lector debe consagrarse menos al desvelamiento de las intenciones del autor que al desciframiento de la “cosa” textual.

Como toda proposición, la que el texto propone debe contar con un sentido y una referencia. Sólo que la referencia es doblemente problemática. De un lado, porque la situación de escritura se caracteriza por suprimir las condiciones de la mostración ostensiva inmediata; y, de otro, porque la proposición procede, no de evento comunicativo dialógico, sino del texto mismo. Aunque la proposición que el texto crea se afina en la observación que el autor hace del mundo, la unión entre sentido y referencia resultante no es un mero reflejo del mundo. Efectivamente, el texto, gracias al efecto de estructuración en obra realizada por el autor, ancla su horizonte de referencia en el mundo real, pero no para manifestarlo en lo que es, como si se tratara de un espejo, sino para re-figurarlo, para imaginarlo, para conjeturarlo, en sus eventuales posibilidades de ser³. El referente que el texto nombra y crea se distancia de la inmediatez del referente designado por el lenguaje ordinario y se abre a nuevas posibilidades semánticas y referenciales. Esa proposición de mundo que el texto actualiza puede o no ser declarada explícitamente. Sí lo es, la tarea de comprensión se orienta en una dirección diferente, bien hacia la intelección de la trama narrativa (en el caso de los textos de ficción), bien hacia la aprehensión de los argumentos con base en los cuales el “autor” pretende crear o aumentar las razones que favorecen la adhesión a la proposición planteada (en el caso de los textos no ficticios).

Sin embargo, ¿y qué hacer cuando la proposición no aparece expresada de un modo inmediato? Ricoeur es claro al respecto: al lector no le queda más opción que corresponder con su actividad descifradora a la labor de estructuración textual acometida por el autor. Y dicha actividad de descodificación se resume en dos actos cognitivos: aventurar una conjetura de sentido y, casi al mismo tiempo, pergeñar un conjunto de argumentos aducido a favor de su conjetura.

Apuntalada en una suerte de intuición adivinatoria, la conjetura debe estar limitada, no por un juego de asociaciones libre, sino por la lógica del mismo texto. En esa medida, conjeturar es todo menos decir aquello

³ “Mediante la ficción, mediante la poesía, se abren en la realidad cotidiana nuevas posibilidades de ser-en-el-mundo-. Ficción y poesía se dirigen al ser, no bajo la modalidad del ser-dado, sino bajo la modalidad del poder-ser” (Ricoeur, *Del texto* 107-08).

que no ha sido sometido al tribunal de la razón textual. Y esta razón no es un agregado de pensamiento que se le impone al texto desde fuera, sino, antes bien, el producto de una morosa permanencia en el interior de su entramado compositivo. A la conjetura se llega luego de que el lector, sirviéndose de las herramientas que, por ejemplo, provee el análisis estructural, está en capacidad de explicar el modo cómo el texto se articula internamente en sus distintos niveles de composición y en las diversas relaciones que dichos niveles sostienen entre sí. Sentido, en este contexto, significa, justamente, establecimiento de las relaciones que los elementos constituyentes del texto establecen unos con otros y que el texto establece con textos que pertenecen a un corpus textual similar. Conforme a este primer acto cognitivo, se trata de leer sin pensar que el texto es tributario de un mundo y de un autor (Ricoeur, *Del texto* 135).

El segundo acto, de orden trascendente (en el sentido de abandonar la inmanencia textual), implica, de una parte, tomar nota de la significación que mejor compete a la referencia liberada por el propio texto, y, de otra, inscribir dicha significación en un orden de discurso inédito que tiene por destino la comunicación a alguien más. A resultas de lo primero, el lector se percibe de nuevas cualidades de la realidad como las desplegadas por el texto y las sopesa en términos de *otras posibilidades de ser* que sólo admiten ser plasmadas de modo discursivo. Como consecuencia del segundo aspecto, el lector explica aquello que ha sido comprendido (el sentido y la referencia constitutiva de la proposición textual) y argumenta por extenso a su favor, en la esperanza de recibir cierta clase de validación por parte de un auditorio general o universal.

La complementariedad de estos dos actos cognitivos desemboca en la noción de apropiación, dado que para Ricoeur no hay comprensión que no sea ella misma una aplicación de lo comprendido a una situación presente. De ahí que forme parte del destino de un verdadero texto

el poder descontextualizarse [de su propia situación, lo cual incluye dejar de lado las preocupaciones por la significación de las intenciones del autor] para que se lo pueda recontextualizar en una nueva situación: cabalmente lo que hace el acto de leer. (Ricoeur, *Del texto* 104)

Leer es, por tanto, dejar valer el texto en su distanciamiento originario. Sólo de esa manera se le puede comprender en lo que detenta de obra estructurada y organizada. Lo que sigue después de ello, a juicio de

Ricoeur, es un

movimiento de apropiación o aplicación del texto a la situación presente del lector... La apropiación es todo lo contrario de la contemporaneidad y de la congenialidad [como lo pretendía la hermenéutica romántica y la hermenéutica diltheyana]; es comprensión por la distancia, a la distancia. (Ibíd. 109)

Sólo si el lector lleva a cabo el proceso de apropiación puede vencer el distanciamiento originario del texto. Volver cercano lo que yacía distanciado, tornar, si no propio, familiar lo que se erguía desconocido, en últimas, “efectuar la referencia”, es la tarea última de un acto de interpretación textual que se apunta en la comprensión. Y esa efectuación no es otra cosa que la puesta en práctica del fin último de la lectura, a saber, “actualizar las potencialidades semánticas de la obra textual” (Ibíd. 141-42), es decir, poder hablar el discurso escrito conforme, no sólo a lo que significan las oraciones que lo constituyen, sino también conforme a la referencia designada por él, así ésta entre en choque con la que es habitual en el lenguaje ordinario. Y quien puede hacerlo no es otro que el lector, que, al hacerlo, puede a su vez comprenderse a sí mismo.

Esta noción de apropiación guarda relación con el concepto de traducción gadameriana.

Comprender un texto o un enunciado textual es, conforme a la interpretación que hace Grondin de Gadamer, “traducir un sentido o ser capaz de traducirlo” (*¿Qué es?* 86).

La definición citada nos reenvía entonces al comienzo de este texto, en el que Gadamer señala la indisociable relación entre la comprensión y el lenguaje. Sólo que ahora se incorpora una idea nueva: la comprensión como *traducción*. Diríase que comprender, en tanto acto similar al de la traducción, equivale a *volver a decir lo dicho* en “un lenguaje que siempre es necesariamente el nuestro” (Ibíd.), o, si se quiere, equivale a decir a otro (o a sí mismo), con palabras “propias”, el sentido aprehendido, el sentido que es producto del discernimiento, pero, eso sí, cuidándose, precaviéndose en lo posible, señala Gadamer, de no confundir las ocurrencias propias, los prejuicios propios, con lo que el enunciado proferido o el texto leído (la “cosa misma”) efectivamente dice (Gadamer,

Verdad I 333). Se comprende para expresar lo comprendido. En esa medida, la expresión de lo comprendido representa una forma de “actualización del lenguaje como discurso”, según el predicamento de Ricoeur arriba mencionado.

Sea cual fuere el caso en que la comprensión acontezca, por tanto, para Gadamer la comprensión es del orden del acontecimiento, es decir, de aquello que irrumpe de un momento a otro sin previo aviso y nos envuelve en un halo de exaltación cognoscitiva (“Lo tengo” o “Comprendo”, decimos sin poder explicar muy bien lo que ha pasado en nuestro cerebro), ella supone verter en expresiones, en enunciados, no la referencia (que siempre es *un más allá signico*), sino el sentido. El sentido comprendido, aquello que nos prende y en lo cual permanecemos temporal o definitivamente, es lo que comunicamos a otros, en la esperanza de restituir acuerdos perturbados o de crearlos cuando no existen.

Sólo que, en Gadamer, la comprensión, como posibilidad o actualización enunciativa de un sentido comprendido, es un proceso dinámico, nunca estático, que exige tener en cuenta dos comportamientos hermenéuticos.

El primero, derivado de retórica antigua, prescribe entender el todo desde la parte y, una vez ejecutado esto, volver al todo para revisar lo comprendido. Dicha regla se conoce con el nombre de *círculo hermenéutico* o “círculo de la comprensión”. Ante ejecuciones de habla o de escritura, la tarea hermenéutica de la comprensión debe ser siempre la misma: no asumir los enunciados como entidades discursivas cerradas sobre sí mismas, sino como entidades que hacen parte de una cadena enunciativa cuyo sentido, al tiempo que arrastra los contenidos significativos emitidos o escritos antes, espera su confirmación en la realización de los enunciados ulteriores.

Se trata de ampliar en círculos concéntricos la unidad de sentido comprendido. La confluencia de todos los detalles en el todo es el criterio para la rectitud de la comprensión. La falta de tal confluencia significa el fracaso de la comprensión. (Gadamer, *Verdad II* 63)

Tal vez por esto, la metáfora es infeliz, pues suscita la idea de que la comprensión se abre y se cierra en un punto determinado. Al cerrarse el círculo la comprensión se completaría. Pero Gadamer es insistente en

señalar que el proceso no ocurre de ese modo. La comprensión nunca se completa del todo, debido a que los textos no agotan completamente su potencia de sentido y al hecho adicional de que el intérprete nunca se baña dos veces en las aguas del mismo texto.

Descrita en estos términos, la comprensión demanda paciencia, técnica, rigor y probidad. Paciencia para permanecer “en la cosa misma”, llámese conversación o texto, sin ceder a la tentación de dejarse ganar por los prejuicios; técnica para ejecutar con maestría unos procedimientos de escucha y lectura aptos para *dejarse hablar* por el otro o para *dejarse interpelar* por el texto leído; rigor para “aventurar hipótesis que habrá de contrastar con las cosas” (Ibíd. 65); y probidad para aceptar que el otro o el texto mismo, en lo que dicen, pueden deslastrar el aferramiento a las propias opiniones y de ese modo multiplicar la apertura de los horizontes de comprensión.

Por lo demás, la comprensión, así sea universal, no se realiza en todos los casos del mismo modo. En primer lugar, en ella hay algo de lo imponderable, de lo irreductible a reglas, que tiene que ver menos con disposiciones subjetivas intencionales que con determinaciones varias ligadas a la lengua que se habla, la comunidad a la que se pertenece, la tradición de la cual se forma parte o a la formación (*Bildung*) que se ha ganado a lo largo de la existencia. Y en segundo lugar, comprendemos de conformidad con el modo como hablamos y escribimos (modos ambos que guardan a su vez estrecha relación con el modo como leemos).

La segunda conducta, que en verdad no es tanto una regla cuanto un complemento hermenéutico, es llamada por Gadamer “anticipo de la compleción o anticipo de perfección” (entendida ésta como la calidad o condición de completo o perfecto). Este elemento consta de dos partes: por un lado, suponemos que al hablar o escribir, el hablante o escritor expresa sus enunciados plenamente, no esquemáticamente, y, por otro, suponemos que lo que expresa es verdad (Gadamer, *Verdad II* 67). La expresión plena de enunciados verdaderos por parte de un hablante o de un escritor garantiza la posibilidad de la atribución de autoridad en el seno de un grupo social determinado. Autoridad, no como expresión de poder, sino como manifestación de buen juicio.

III

Concluyendo, cuesta trabajo concebir un ser humano cuya existencia vuelva la espalda a la comprensión. Gracias a los trabajos iniciales de

Heidegger hemos tomado conciencia de que nuestras vidas están signadas originariamente por un comprender primario en virtud del cual habitamos el mundo de un modo circunspectivamente comprensivo. Nada en esta fase sugiere que el ser humano empieza a vivir su vida fáctica analizando el mundo en el cual se desenvuelve, a semejanza de un observador tocado por preocupaciones conceptuales, por no decir vitales, que da por sentado que la vida, encerrada dentro de sus propios límites y con unos atributos especiales, existe más allá de sí en un presunto “afuera” soberano que pudiera convertirse en objeto de investigación.

¿Qué ocurre cuando los objetos a la mano se nos resisten, cuando la familiaridad que esperamos encontrar en el mundo en el cual somos arrojados interrumpe su comparecencia habitual y produce un alto en nuestra comprensión circunspectiva de ese mismo mundo, o cuando, en una palabra, notamos que las cosas no son como debieran ser?

Contamos con dos alternativas: primera, tendemos a restituir el equilibrio que las cosas a la mano han alterado, reemplazando en el caso de los objetos, uno por otro, y propiciando de ese modo la continuación de nuestro comportamiento habitual, pues en últimas lo que pretendemos es que las cosas *vuelvan a tener sentido* para nosotros (en el entendido no articulado lingüísticamente de que “sentido es aquello [el horizonte] en lo que se mueve la comprensibilidad de algo”) (Heidegger, *Ser y Tiempo* 175); segunda: ponemos en marcha, como proceso derivado del comprender originario, del “cómo hermenéutico”, si no el conocimiento teórico, la articulación enunciativa de aquello que causa la incomprensión. En esta medida, “comienzo, nos dice de nuevo Bauman,

a plantear las preguntas cuando las cosas cesan de formar parte del mundo vital al cual estoy familiarizado y, por ende, se presentan ante mí como una cosa-en-sí-misma, un objeto extraño que debe ser considerado a través del análisis teórico (154).

Tomo distancia de mi propia situación y considero el objeto, al que ya no veo como ser-a-la-mano solamente, sino también como pudiendo no estar allí o como pudiendo ser de un modo diferente. Si la primera alternativa es ante-predicativa, lo que no significa a-lingüística, la segunda es enunciativa o si se prefiere, discursiva, ya que ahora se trata de tematizar, de exponer expresivamente, aquello que se resiste a la comprensión.

Gadamer y Ricoeur, al meditar hermenéuticamente sobre la actividad de la comprensión, coinciden en afirmar que el lenguaje se actualiza discursivamente en el diálogo, y que en éste todo fenómeno de incompreensión, comprensión a medias o malentendido, como consecuencia de la inexistencia o alteración de los acuerdos sociales o de la polisemia constitutiva del mismo lenguaje, encuentra en el modelo de las preguntas y respuestas un espacio vivo para procurar restablecer o instaurar unos mínimos racionales de conversación o unos mínimos de univocidad, necesarios (aunque no suficientes) para superar la parálisis del acto comunicativo o la ambigüedad de lo dicho.

Y si Ricoeur, en lo que atañe a los textos, es explícito en aseverar que la tarea del lector consiste en moverse, conforme a un amplio "arco hermenéutico", de la explicación a la comprensión y de ésta aquélla, con el propósito de coronar interpretativamente un discurso fijado por escrito, Gadamer propone operar no sólo con los principios de anticipación de compleción y de círculo de comprensión, sino también con una conciencia dramática del lenguaje que no silencia las determinaciones de los prejuicios con base en los cuales encaramos las producciones del espíritu que hacen parte de la historia de la humanidad.

En suma, tal vez más temprano o más tarde, todo depende de las *figuraciones de nuestras conciencias*, comprendamos que la comunicación, cuando se asienta en los dominios operativos del habla y la escucha, o de la lectura y la escritura, antes que una labor mecánica de transmisión de información (regulada por secuencias fijas de instrucciones programadas con antelación por alguien distinto de nosotros mismos), es un amplio campo de acciones verbales recursivas en el que buscamos llegar a acuerdos mínimos con nosotros mismos y con los otros, a través de tanteos, balbuceos o bregas permanentes, sobre aquellos asuntos de nuestras vidas públicas, privadas, imaginarias o referidas que más nos afectan y que más afectan a los demás.

REFERENCIAS BIBLIOGRÁFICAS

Bauman, Zygmunt. *La hermenéutica y las ciencias sociales*. Buenos Aires: Nueva Visión, 2007. Impreso.

Echeverría, Rafael. "Sobre el lenguaje humano". *Ontología del lenguaje*. España: Dolmen Ediciones, 2002. Impreso.

Gadamer, Hans-Georg. *Verdad y Método II*. Salamanca: Sígueme, 1994. Impreso.

---. *Verdad y Método I*. Salamanca: Sígueme, 1996. Impreso.

Grondin, Jean. *Introducción a Gadamer*. España: Herder, 2003. Impreso.

---. *¿Qué es la hermenéutica?* Barcelona: Herder, 2008. Impreso.

---. *El legado de la hermenéutica*. Santiago de Cali: Programa Editorial Universidad del Valle, 2009. Impreso.

Heidegger, Martin. *Ser y Tiempo*. Madrid: Trotta, 2006. Impreso.

---. *Ontología: hermenéutica de la facticidad*. Madrid: Alianza Editorial, 2007. Impreso.

Husserl, Edmund. *Investigaciones lógicas I*. Madrid: Altaza, 1995. Impreso.

Ricoeur, Paul. *Teoría de la interpretación. Discurso y excedente de sentido*. México: Siglo XXI, 2003. Impreso.

---. *Del texto a la acción. Ensayos de hermenéutica II*. México: Fondo de Cultura Económica, 2002. Impreso.

FENOMENOLOGIA HERMENÊUTICA, CIRCULARIDADE E DESOCULTAMENTO

PHENOMENOLOGY HERMENEUTICS, CIRCULARITY AND UNCONCEALMENT

CEZAR LUÍS SEIBT

Universidade Federal do Pará, Brasil. celuse@ufpa.br

RECIBIDO EL 3 DE MARZO DE 2012 Y APROBADO EL 6 DE ABRIL DE 2012

RESUMO ABSTRACT

O texto mostra como Heidegger elabora sua proposta de superação da metafísica, da abertura para o acontecimento contra a objetificação possibilitada pela tradição, passando por uma nova visão da fenomenologia, ligada à hermenêutica. Enfatiza a questão dos pressupostos, considerando-os como condições prévias para o acontecimento do mundo e do reconhecimento dos entes. Faz o conhecimento entrar autenticamente no círculo hermenêutico e reconhece suas potencialidades. No processo de desocultamento dos entes, as condições prévias, o horizonte do ser, mantêm-se ocultos. Tanto os entes como as condições prévias do acontecimento dos entes se dá dentro do círculo hermenêutico.

This paper shows how Heidegger develops his proposal to overcome metaphysics, the opening for the event against the objectification possible by tradition, through a new vision of phenomenology linked to hermeneutics. It emphasizes the issue of presuppositions, regarding them as preconditions for the event of the world and the recognition of the entities. It gets the knowledge authentically in the hermeneutic circle and recognizes its potential. In the process of unconcealment of beings and preconditions, the horizon of being remains hidden. Both entities and preconditions of the event take place within the hermeneutic circle.

PALAVRAS CHAVE KEY WORDS

Círculo hermenêutico, conhecimento, fenomenologia hermenêutica, ocultamento.

Hermeneutic circle, knowledge, hermeneutic phenomenology, concealment.

Considerações introdutórias: da fenomenologia a fenomenologia hermenêutica

Na entrevista de Heidegger para *Der Spiegel*, que saiu com o título: “Meu caminho para a fenomenologia”, Heidegger avalia e mostra, de forma geral, o modo como foi aos poucos se diferenciando e distanciando da fenomenologia de Husserl, tendo como pano de fundo sua leitura de Aristóteles. Percebe que a atitude básica da fenomenologia é aquela que caracterizava já o modo grego de filosofar. Um retorno às coisas mesmas tinha a ver com desocultamento, desvelamento, o que Heidegger lê na palavra *Alétheia*. Nas palavras do próprio pensador:

[...] o que para a fenomenologia dos atos conscientes se realiza como o automostrar-se dos fenômenos é pensado mais originariamente por Aristóteles e por todo pensamento e existência dos gregos como *Alétheia*, como o desvelamento do que se pre-senta, seu desocultamento e seu mostrar-se. Aquilo que as Investigações redescobriram como a atitude básica do pensamento revela-se como o traço fundamental do pensamento grego, quando não da filosofia como tal. (*Pensadores* 300)

Nesse encontro com Aristóteles e a filosofia grega, Heidegger confronta-se com o problema do sentido do ser que, no início parece ter certa vitalidade e ser fundamental para os demais problemas filosóficos, mas que aos poucos emudece a partir de uma naturalização de um dos modos possíveis de compreensão.

A determinação do tempo como presença, fixa também o modo de compreensão do ser e, com isso, Heidegger põe em dúvida a capacidade de experimentar a coisa mesma a partir da ontologia tradicional. A fenomenologia parece oferecer uma saída, mas não a partir do modelo da reflexividade. É preciso perguntar pelo ser do desvelar, do apresentar. O problema é saber qual é a coisa mesma a ser investigada, se a consciência, como no caso da fenomenologia de Husserl, ou o ser do ente no seu desvelar-se. Segundo Heidegger, quanto mais se ocupava do aspecto fundamental da filosofia grega, que é a ideia de desocultamento ou *Alétheia*,

[...] tanto mais insistente tornava-se a questão: de onde e de que maneira se determina aquilo que, de acordo com o princípio da fenomenologia, deve ser experimentado como

‘a coisa mesma’? É ela a consciência e sua objetividade, ou é o ser do ente em seu desvelamento e ocultação? Deste modo fui levado ao caminho da questão do ser, iluminado pela atitude fenomenológica, sempre de novo e cada vez de maneira diferente, inquietado pelas questões emanadas da dissertação de Brentano. O caminho do questionamento tornou-se, porém, bem mais longo do que presumia. Exigia muitas paradas, desvios e sendas perdidas. (Heidegger, *Pensadores* 300)

Heidegger pergunta pelo sentido do ser a partir da questão da coisa mesma. Nesse perguntar e no confronto com o pensamento grego, (re)-descobre o caráter de ocultamento no descobrimento. No aparecer do ente, na sua manifestação, acontece um encobrimento. No desvelar há o velar; quando o ente aparece, desaparece o fundo, o horizonte, dentro do qual o ente é possível. Desaparece o próprio ser-no-mundo. Stein afirma que: “o fator determinante e individualizador do método fenomenológico é a descoberta que Heidegger fez de que existe um primado da tendência para o encobrimento” (89-90). A fenomenologia lida, portanto, não principalmente com o que se dá, mas, sobretudo, com o que se oculta no dar-se.

Para que se possa pensar esse caráter de desvelamento e ocultamento, ou seja, a revelação do ser no ente enquanto um duplo movimento, não se pode ficar preso ao ente que se manifesta na abertura. Há que retornar para o modo de ser no qual e através do qual se dá o fenômeno da abertura, ao *Dasein*, portanto. Uma hermenêutica desse ente irá:

[...] pôr a nu os modos de dissimulação em que primeiramente, e o mais das vezes, se situa o próprio ser-aí, na sua cotidianidade. Heidegger descobre o ser-aí no movimento de fuga de si mesmo, numa tentativa de não se assumir na sua totalidade, como preocupação, que se articula em existência, faticidade e decaída ou ser-adiante-de-si, já-ser-em e junto-dos-entes. O ser-aí se vela para si mesmo, encobre suas possibilidades e assim barra a possibilidade de uma revelação do ser. A atitude do filósofo para contornar a fuga do ser-aí de si mesmo, é partir da análise da cotidianidade e descobrir nela o homem no movimento de fuga. (Heidegger, *Pensadores* 89-90)

Mas precisamos deixar mais clara à questão da ‘coisa mesma’. Quais são as coisas? A que coisas mesmas? Heidegger mesmo responde:

Por um lado, às coisas mesmas no sentido de investigar demonstrando as coisas com os pés postos no solo; por outro, em primeiro lugar recuperar e assegurar esse solo (o olhar; exigência de colocar o solo a descoberto). A segunda delas é a exigência de colocar fundamento e por isso inclui a primeira. (Heidegger, *Pensadores* 104)

Às 'coisas mesmas' não significa a descrição de objetos particulares e específicos, mas a investigação que põem os pés no solo, e que para colocar os pés no solo precisa antes recuperá-lo. A descrição dos objetos particulares e específicos é tarefa das ciências ônticas, enquanto a filosofia tem a função de demonstrar o solo onde os pés estão apoiados quando alguém se ocupa com objetos.

A filosofia tem um caráter ontológico, de investigação do fundamento, mesmo quando esse fundamento seja mobilidade (*Bewegtheit*). Por isso,

a fenomenologia, enquanto disciplina descritiva, é uma ciência prévia (*Vorwissenschaft*) às disciplinas tradicionais em filosofia, nas quais se discutem as questões concretas. (Ibíd. 107)

Diferentemente das outras 'ciências', a fenomenologia,

não diz nada sobre o conteúdo de coisa do objeto de interesse da ciência, mas precisamente diz acerca do *como*, a maneira de tratar algo e como deve ser tratado em trabalho de investigação. (Ibid. 115)

Ela é uma atitude fundamental que "nada diz acerca do ser do objeto em questão, mas nomeia seu modo de ocorrência" (Ibid. 115).

O ocultamento não pode ser entendido como se por detrás do fenômeno houvesse algo escondido, como se o fenômeno fosse algo atrás do qual houvesse algo a ser descoberto. Não há outro ente ou algo do ente escondido, que pertence ao nível ôntico, mas o que se esconde é do nível ontológico, da condição de possibilidade. O próprio Heidegger diz que:

não se pode perguntar por nenhum detrás, posto que o que se dá é justamente algo em si mesmo. O que sim pode acontecer é que o que se mostra em si mesmo e se tem de demonstrar esteja encoberto. (Heidegger, *Prolegómenos* 116)

O trabalho da filosofia consiste em “desmontar os encobrimentos”, em “fazer ver colocando em descoberto” (Heidegger, *Prolegómenos* 116).

O modo tradicional de entender a realidade está baseado nos entes que estão manifestos. O que se torna agora problema da fenomenologia é o horizonte que possibilita a manifestação, o que não se manifesta na manifestação dos entes, mas, ao contrário, se oculta nesta manifestação. Fenomenologia em sentido fenomenológico, como Heidegger agora pensa, não é descrição de entes que se manifestam na abertura, mas a descrição da própria abertura, pois esta se oculta com o manifestar dos entes. O encobrimento há de tornar-se assunto da fenomenologia.

Para que se possa falar de algo como encobrimento, esse não pode ter se realizado em totalidade, precisa haver ainda algo presente, mesmo que em forma de aparência. Na aparência e também no mascaramento os “fenômenos originariamente vistos se mantêm desarraigados, arrancados do seu solo e sem entender-se em sua procedência efetiva” (Ibid. 116-17).

Essa é a crítica de Heidegger aos conceitos teóricos que se desgrudam do solo pré-teórico onde se originaram e acabam sendo assumidos irrefletidamente e de forma dogmática. O dogmatismo acontece se for feito um corte entre o conhecimento (o conhecido) e as condições prévias e pré-teóricas que tornam possível o comportamento teórico. A coisa originariamente experimentada em sua vigor e efetividade podem

[...] ser transmitida de maneira vazia, segundo um modo de entender já predeterminado, perdendo o enraizamento no solo e tornando-se voz que flutua no ar. Essa possibilidade de petrificação do originariamente obtido e acreditado está implícita no próprio trabalho concreto da fenomenologia.

[...] junto com a possibilidade do descobrir radical leva em si o perigo correspondente de fixar-se nos seus êxitos. (Ibid. 116)

Por isso, o trabalho fenomenológico é um “ater-se crítico, em sentido positivo, consigo mesmo [...] para deixar-se guiar e determinar pelo que significa em cada caso o fenomenologicamente dado da intencionalidade” (Ibid. 117-18).

Não há nenhuma comodidade neste trabalho. Há um trabalho a ser feito consigo mesmo, de desconstrução ativa de pressupostos sem, contudo, eliminar a dimensão do pressuposto. Dito de outra maneira há que lembrar o esquecimento, ou desocultar o ocultamento. Isso exige um esforço que Heidegger resume na dupla tarefa de fazer uma analítica do *Dasein* e na destruição da metafísica.

Ocupar-se com o 'retorno às coisas mesmas' conduz para a questão básica do sentido do ser que toma como ponto de partida o lugar onde o ser é compreendido.

Esse lugar é um ente que precisa, ele mesmo, por ser o ente que tem uma compreensão do ser, autointerpretar-se primeiramente para libertar-se dos esquecimentos de que é vítima em função da sua ocupação com os entes que se dão na abertura. Essa libertação é um desocultamento do seu modo fundamental de ser, que tem implicações fortes na posterior elaboração da questão fenomenológica básica do sentido do ser, uma questão que:

nenhuma ontologia conseguiu colocar e para a qual, no entanto, de alguma maneira, de modo fundado ou infundado, constantemente se-lhe pressupõe uma resposta e se faz uso dela. (Heidegger, *Prolegómenos* 122)

E, como diz Dartigues,

a fenomenologia hermenêutica deverá, pois, decifrar o sentido do texto da existência, esse sentido que precisamente se dissimula na manifestação do dado. (115)

Alcançar esse horizonte que foi ocultado e ignorado na investigação tradicional e que em Husserl aparece no fato de deixar intocado o ser da intencionalidade é fundamental para que se possa ver que "as chamadas atuações lógicas do pensar ou do conhecer teórico, objetivador, não constituem senão uma parte concreta e reduzida" (Heidegger, *Prolegómenos* 106).

O esforço se concentra na tentativa de mostrar que o conformar-se às coisas mesmas é um "apreender originário simples e não alguma espécie de construção experimental de subestruturas" (Ibid. 106).

Fenomenologia hermenêutica e a circularidade da compreensão

Heidegger indicou pontos centrais do caminho que havia percorrido:

companheiro na busca foi o jovem Lutero e o modelo foi Aristóteles, a quem aquele odiava. Impulso me deu Kierkegaard e os olhos foi Husserl quem me implantou. (Heidegger, *Ontologie* 5)

Diversas fontes se juntam no pensamento do jovem filósofo, que vão desde a ontologia clássica até a crítica da mesma, que se realiza no apreço pela singularidade, pela existência particular.

Essa problemática da universalidade, do fundamento, se faz presente no seu pensamento, mas profundamente marcado pela existência, pela faticidade. A questão ontológica é marcada pela realidade fática. O ser não é mais um ser teórico, uma abstração, mas um ser que é compreendido no horizonte de um determinado ente. Com essa junção, nem ser, nem o ente *Dasein* podem ser compreendidos como antes. Há uma nova visão aqui, um novo 'paradigma'.

Heidegger não quer repetir o 'equivoco' de Husserl e vai procurar realizar uma fenomenologia hermenêutica, ou seja, uma fenomenologia do ser que tem como ponto de partida o ser do *Dasein*. E, como diz Gethmann, a filosofia de Heidegger "terá como tema a ontologia e como método a fenomenologia" (21) ou, como diz o próprio Heidegger "nós denominamos o método da ontologia de *fenomenologia*" (*Die Grundprobleme* 466).

Stein assim explica o sentido da hermenêutica em Heidegger:

hermenêutica não significa, em *Ser e Tempo*, nem teoria da arte de interpretar, nem a própria interpretação, antes a tentativa de, primeiramente, determinar a essência da interpretação, a partir do hermenêutico. (187)

O *Dasein* é ele mesmo, hermenêutico. Ele existe interpretando, enredado num círculo que finca suas raízes na faticidade e, a partir dela, avança sem poder se libertar deste pressuposto prático.

Com a caracterização da fenomenologia do *Dasein* enquanto hermenêutica marca-se a diferença em relação à fenomenologia de Husserl. Mas é no próprio confronto com o pensar de Husserl que emerge a diferença. Fenomenologia enquanto hermenêutica coloca a base do projeto na existência fática do *Dasein* enquanto ente a partir de onde há compreensibilidade e, portanto, ser.

A hermenêutica fenomenológica deverá tornar-se pesquisa radical, a partir da faticidade. Tal fenomenologia será hermenêutica pelo fato de que é uma autointerpretação da faticidade. Stein indica isso quando diz que:

[...] o método fenomenológico visa ao redimensionamento da questão do ser, não numa abstrata teoria do ser, nem numa pesquisa historiográfica de questões ontológicas, mas numa imediata proximidade com a práxis humana, como existência e faticidade. A linguagem -o sentido, a significação- não é analisada num sistema fechado de referências, mas ao nível da historicidade.

[...] O método fenomenológico, enquanto método hermenêutico-lingüístico, não se desliga da existência concreta, nem da carga pré-ontológica que na existência já vem sempre antecipada. É isto que lhe dá como característica uma inelutável circularidade. (Heidegger, *Pensadores* 88)

Para fugir da determinação abstrata do ser que leva a uma concepção também abstrata do ser humano, Heidegger se dirige para a faticidade, para a existência mesma do ente que já sempre compreende o ser.

A partir dele, numa autointerpretação como ponto de partida, se avança para as outras questões tradicionais que perderam suas raízes ao longo da história. Isso quer dizer, retomar as raízes temporais, históricas do compreender humano só pode dar-se a partir do lugar onde a compreensão se realiza. Esse lugar onde ela acontece não é nenhum ente objetivo que possa ser determinado com base nas categorias tradicionais do tempo presente, mas passam por uma hermenêutica, uma interpretação que não tem como evitar a circularidade.

Dentro da circularidade acontece a compreensão do *Dasein* e nessa circularidade ela permanece, na maior parte das vezes e em geral, de forma inautêntica e imprópria. Ela pode conquistar sua autenticidade

e propriedade sem sair do círculo ou ancorando-se em algum porto seguro, mas assumindo a própria circularidade, a condição de ser-no-mundo como sua condição de ser. Isso implica também uma nova análise da questão do tempo.

Vida não é mais um fenômeno que pode ser apreendido objetivamente, mas é '*Da-sein*': é existência enquanto fenômeno temporal e abertura. Mas isso não é tão simples se recordarmos o caráter duplo de encobrimento e desencobrimento que a atenção ao fenômeno nos manifesta. É o ser do *Dasein* que em primeiro lugar é preciso enfrentar e tornar acessível no seu aparecer marcado pelo ocultamento. Nisso vemos o esforço de Heidegger em *Ser e Tempo* ao tomar como ponto de partida a descrição do *Dasein* em sua cotidianidade. Aquilo que antes dizíamos da *Alétheia* grega tem seu impacto na autocompreensão do ente que compreende o ser.

Somente através da fenomenologia, agora tomando consciência da ambigüidade que ocorre no fenômeno da manifestação, ou seja, através da radicalização da fenomenologia de Husserl, Heidegger estabelece um ponto de partida para o trabalho filosófico: uma hermenêutica do *Dasein*. Dessa forma,

[...] enquanto utilizada para a analítica da faticidade e da existência, a fenomenologia se torna hermenêutica, passa a se movimentar num círculo hermenêutico. Esta circularidade, que não é apenas característica da compreensão, mas através dela, do próprio ser-aí, também apresenta uma ambigüidade que acompanha toda a obra de Heidegger. Pelo método fenomenológico se desvendou esta circularidade, que passa, por sua vez, a possibilitar uma verdadeira penetração na fenomenologia. (Heidegger, *Pensadores* 94)

No texto que resultou do seminário do semestre de verão de 1923, intitulado *Ontologia (Hermenêutica da faticidade)* o autor começa a tratar de forma explícita do caráter hermenêutico da fenomenologia e caracteriza a ontologia como fenomenologia. Isso que aí se inicia continua se desenvolvendo até resultar na analítica existencial do *Dasein* em *Ser e Tempo*.

Na medida em que a fenomenologia é hermenêutica, indica-se que a fenomenologia não é pura. Ela

está carregada de pressupostos do mesmo modo que todo conhecimento humano, e o filosófico não consiste em eliminar os pressupostos a qualquer preço por meio de argumentos, mas reconhecê-los e ajustar a investigação positiva e objetivamente conforme eles. (Heidegger, *Ontologie* 223-24)

Heidegger também reconhece que há pressupostos que uma investigação filosófica nunca alcança que sempre permanecem às suas costas e que, portanto, não permitem transparência total. Na compreensão nem tudo pode ser exposto ou demonstrado de forma plena e definitiva. Há sempre um resto que lhe escapa.

Olhando para o caminho de Heidegger é possível constatar que no princípio a “intencionalidade está no centro da investigação e orienta-se cada vez mais claramente em direção à ‘vida em si’, à vida fática” (Ruckteschell 46). Distancia-se do modelo metafísico para aproximar-se da vida fática.

No curso *Ontologia (Hermenêutica da faticidade)*, o parágrafo três explicita o que o autor compreende por hermenêutica: é a interpretação que a faticidade realiza de si mesma. O tema da hermenêutica é o próprio existir, o próprio *Dasein*, exatamente pelo fato de o tema ele mesmo ser hermenêutico.

E o ponto de partida não poderá ser nenhuma idealização do ser humano, mas o seu modo concreto de ser, a medianidade e cotidianidade na qual o *Dasein* concreto e efetivo se realiza.

Parte, por isso, de uma interpretação vigente, prévia; fala a partir do já interpretado para o já interpretado (Heidegger, *Ontologie* 18) sem pretensão de alcançar um ponto externo para avaliar a interpretação. Isso porque a existência “não é nunca ‘objeto’, mas ser; existe, está aqui somente enquanto ‘é’ o viver de cada momento” (Ibid. 19).

Faticidade, como o próprio Heidegger diz, é o “nome que damos ao caráter de ser do nosso próprio existir. Mais exatamente a expressão significa: esse existir em cada ocasião” (Heidegger, *Ontologie* 7).

Faticidade não pode ser nenhuma abstração ou o resultado de algum silogismo. Não é o resultado de nenhum esforço teórico, mas a constatação

imediate do existir ocasional. A faticidade precisa interpretar a si mesma e isso Heidegger chama de hermenêutica. E essa interpretação é possível, “porque os modos fundamentais não são livremente oscilantes, pois o *Dasein* existe faticamente” (Marx 98).

O *Dasein* não é nenhuma essência, não possui determinação prévia que o caracterize como *Dasein*, mas, por outro lado, também não é qualquer coisa que flutue livremente por aí. O *Dasein* é situado, existe sempre faticamente, numa certa transcendência horizontal, imanente.

A fenomenologia hermenêutica, muito mais do que uma construção teórica, é a constatação da finitude, historicidade, faticidade do ser humano. Ela conquista para ele um novo modo de estar no mundo e no meio dos entes. Parte das condições históricas e concretas inultrapassáveis e incontornáveis nas quais o ser humano existe, já sempre numa abertura compreensiva, no ambiente lingüístico que o abre para o sentido das coisas e do mundo. Existe na condição de intérprete e essa compreensão se constitui se abre, na própria relação, num círculo de interpretação que institui o ser, o sentido das coisas, do mundo e autocompreensão do *Dasein*.

Essa abertura compreensiva (acontecimento do ser) antecede e é condição de possibilidade do conhecimento das ciências particulares diversas. Na impossibilidade de sair da abertura compreensiva sem, com isso, fechar-se para o encontro com qualquer ente, o *Dasein* encontra-se num círculo interpretativo. Ser *Dasein* é já sempre interpretar, encontrar-se aberto para o acontecimento do mundo. Sem nenhuma referência externa à própria abertura, não tem como superar a condição hermenêutica.

REFERÊNCIAS BIBLIOGRÁFICAS

Dartigues, André. *O que é a fenomenologia?* São Paulo: Centauro, 2005. Impresso.

Gethmann, Carl Friedrich. *Dasein: Erkennen und Handeln-Heidegger im phänomenologischen Kontext*. Berlin/New York: Walter de Gruyter, 1993. Impresso.

Heidegger, Martin. *Die Grundprobleme der Phänomenologie*. Frankfurt/Main: Vittorio Klostermann, 1975. Impresso.

- . *Heidegger (Pensadores)*. São Paulo: Abril Cultural, 1983. Impreso.
- . *Phänomenologische Interpretationen zu Aristoteles*. Frankfurt/Main: Vittorio Klostermann, 1985. Impreso.
- . *Ontologie (Hermeneutik der Faktizität)*. Frankfurt/Main: Vittorio Klostermann, 1995. Impreso.
- . Impreso. *Lógica-la pregunta por la verdade*. Madrid: Alianza Editorial, 2004. Impreso.
- . *Prolegómenos para una historia del concepto de tiempo*. Madrid: Alianza Editorial, 2007. Impreso.
- Marx, Werner. *Heidegger und die tradition*. Stuttgart: W. Kohlhammer, 1961. Impreso.
- Ruckteschell, Peter von. *Von der Ursprungswissenschaft zur Fundamentalontologie. Die Intentionalität als Leitstruktur im frühen Denken Heideggers*. Freiburg: Universidade de Freiburg, 1998. Impreso.
- Stein, Ernildo. *Compreensão e Finitude-estrutura e movimento da interrogação heideggeriana*. Ijuí: Unijuí, 2001. Impreso.

LA UNIÓN ALMA-CUERPO EN TOMÁS DE AQUINO UN DIÁLOGO CON A. KENNY

THE UNION BODY-SOUL IN THOMAS AQUINAS IN DIALOGUE WITH A. KENNY

PATRICIA MOYA

Universidad de los Andes, Chile. pmoya@uandes.cl

RECIBIDO EL 26 DE ENERO DE 2012 Y APROBADO EL 6 DE ABRIL DE 2012

RESUMEN

En este artículo, se analizan algunos de los argumentos que Tomás de Aquino expone en la primera parte de la *Suma Teológica*, concretamente: el artículo 4 de la cuestión 75 y el 1° de la cuestión 76; en los cuales el Aquinate trata la unión del alma con el cuerpo. El objetivo es contrastar el análisis de tipo analítico que realiza el filósofo inglés A. Kenny en dos de los capítulos (11 y 12), de su obra: *Aquinas on mind*, con los argumentos tomasianos basados principalmente en principios metafísicos. Con base en lo anterior, se busca dar respuesta a las objeciones que Kenny plantea y a la vez esclarecer la perspectiva desde la cual, Tomás de Aquino sostiene la unidad de la persona humana.

PALABRAS CLAVE

Alma-cuerpo, A. Kenny, entendimiento, forma, Tomás de Aquino.

ABSTRACT

In this paper, some of the arguments advanced by Thomas Aquinas in the first part of *Summa Theologiae* are analyzed. I focus specifically in I, q. 75 a. 4 and q. 76 a. 1, in which Aquinas deals with the union between body and soul. My aim is to contrast the analytical interpretation of A. Kenny in *Aquinas on mind*, chapters 11 and 12, with Aquinas' arguments based mainly on metaphysical principles. Based on that, I try to respond to Kenny's objections and, at the same time, to clarify Aquinas' doctrine about the unity of the human person.

KEY WORDS

Body-soul, A. Kenny, understanding, form, Thomas Aquinas.

Introducción

Es sorprendente la argumentación que realiza Tomás de Aquino en el capítulo 79 del libro IV de *Contra Gentiles*, al defender la conveniencia de la resurrección del cuerpo al final de los tiempos. En este texto, afirma que la separación del alma y del cuerpo provocada por la muerte, es *contra naturam* y que una situación tal, no puede ser perpetua de ahí que sea preciso que el alma que permanece eternamente se una nuevamente al cuerpo en cuanto éste, sea resucitado por el poder de Dios. Sin olvidar que en la mente del Aquinate, la resurrección del cuerpo es considerada siempre como un hecho sobrenatural y que esta nueva unión con el cuerpo se realizará sólo por la acción divina, no deja de llamar la atención, la fuerza argumentativa con la que se muestra la conveniencia de dicha resurrección, justamente por la intrínseca unidad del alma y el cuerpo en la persona humana.

A. Kenny, ha tenido el mérito de reconocer el aporte que el pensamiento tomasiano puede llegar a significar para las discusiones contemporáneas de la filosofía de la mente en particular, en lo que se refiere a la unión del alma y el cuerpo tema que desde Descartes se ha hecho particularmente problemático para la filosofía moderna y contemporánea. Su obra *Aquinas on mind*, nos presenta el análisis desde las cuestiones 75, hasta la 89, de la *Suma Teológica*, cuestiones que forman parte de: “El Tratado del Hombre” en el cual el Aquinate, analiza la unidad corpóreo-espiritual de la persona.

En este artículo, me propongo presentar algunos aspectos de este tema y establecer un diálogo entre la doctrina tomasiana y algunas reflexiones que Kenny realiza, particularmente, en los capítulos XI y XII de la obra anteriormente mencionada.

Me serviré del análisis de dos de los artículos de la *Suma Teológica*, que son fundamentales para tratar diversas cuestiones que se suscitan en torno a la comprensión de la persona humana que tiene Tomás de Aquino y en especial en torno a la unidad alma-cuerpo, tema central de este estudio. Se trata, del artículo 4, de la cuestión 75 y el 1º, de la cuestión 76. Son dos textos conocidos y frecuentemente comentados, pero de los cuales me quiero valer, tanto para exponer la doctrina tomasiana, como para confrontarlos con los análisis que realiza Kenny.

I

La pregunta que dirige la argumentación del artículo 4, c. 75, es: “si el alma es o no es el hombre”. Tomás de Aquino, examina dos alternativas: i) afirmar que el hombre es el alma, pero que este hombre, no es el alma, sino un compuesto del alma y del cuerpo; y ii) decir: *esta alma es este hombre*.

Si se examina la primera alternativa que propone una diferenciación sustancial entre la consideración del hombre como especie y la del hombre como individuo se concluye que no puede admitirse tal distinción, porque no es verdad, como sostuvo Averroes que sólo la forma pertenezca a la especie, mientras que la materia es parte del individuo y no de la especie (80v-6v)¹.

La doctrina de Tomás de Aquino al respecto es que a la naturaleza de la especie, le pertenece lo que expresa la definición y en las cosas naturales, la definición no expresa sólo la forma, sino también, la materia y la forma. La materia que es parte de la especie en las cosas naturales es la materia común, no la materia signada que es principio de individuación:

Pues, así como es propio de la razón de *este hombre* que lo sea a partir de esta alma, carne y huesos, así también es propio de la razón de hombre que lo sea a partir del alma, carne y huesos. Ya que es necesario que la sustancia de la especie tenga lo propio y común de todos los individuos contenidos en dicha especie. (Aquino c. 75, a. 4)²

La segunda alternativa, sólo sería viable si las operaciones del alma sensitiva se realizaran sin el cuerpo, porque de este modo todas las operaciones atribuidas al hombre le corresponderían tan solo al alma. En tal caso, el hombre se definiría exclusivamente por el alma. Pero esta posibilidad, no es admisible, porque el sentir, no es una operación exclusiva del alma, sino que se realiza en estrecha relación con el cuerpo.

¹ Se puede ver más detalladamente la argumentación de Averroes, tal como es expuesta por Cayetano en: *Comentarios al Del ente y la esencia*.

² Ver *El ser y la esencia*, cap. II, donde expone, que en las sustancias compuestas, la esencia, no es sólo la materia o la forma, sino ambos principios. Para resolver la dificultad, que al comprender la materia, la esencia, sólo pueda ser considerada particularmente, lo que es opuesto a su misma noción, explica el Aquinate, que la materia que es principio de la esencia, es la materia común, porque la que individualiza al sujeto, es la materia signada (ver también cap. V). Esta doctrina, es propia de Tomás de Aquino, debido a que en este punto, discrepa claramente de Aristóteles (cf. *Metafísica* VII 10-11).

En el artículo 3° de la misma cuestión, Tomás de Aquino adopta sin variación la postura aristotélica, según la cual el entender se realiza sin órgano corporal, a diferencia del sentir, que realiza sus operaciones contando con los órganos corporales correspondientes y a partir de alguna inmutación corporal (Aristóteles, *De anima* 3 4 429a 24 429b 10).

El interlocutor que Tomás tiene en ambos artículos es Platón, quien le atribuye un principio incorpóreo al sentir, de manera que el alma es principio del sentir y el cuerpo es sólo su instrumento.

En estos argumentos, advertimos la idea que guiará toda la doctrina del Aquinate, acerca de la naturaleza del hombre; éste es una unidad de alma y cuerpo que no puede ser desarticulada bajo ningún aspecto: ni desde el punto de vista metafísico, debido a que su naturaleza sólo se comprende adecuadamente, desde la perspectiva de la unidad sustancial alma-cuerpo, ni tampoco, desde el punto de vista de la operación de sus facultades, porque la comprensión del obrar humano requiere de esta unidad, sin la cual se suscitan dificultades serias como las que se derivan de los planteamientos de Averroes en la explicación de la operación intelectual (Aquino, *Suma Teológica* c. 79 a. 4-5).

Refuerza esta argumentación, la respuesta a la segunda objeción, la cual plantea que el alma es hipótesis o persona y que, por tanto, el alma es el hombre, debido a que la persona humana es el hombre. La respuesta es contundente y la cito completa:

No toda sustancia particular es hipótesis o persona, sino aquella que tiene toda la naturaleza de la especie. De ahí que la mano o el pie no puedan ser llamados hipótesis o persona. De forma parecida, tampoco el alma, ya que es sólo parte de la especie humana. (Ibíd. c. 75 a. 4-ad 2)

Puede resultar chocante la afirmación que el alma es *parte* de la especie humana, pero como justamente apunta Kenny lo que se está tratando de resaltar es la unidad sustancial de dos principios, ninguno de los cuales por separado constituye realmente a la persona humana.

En este sentido, el alma es *parte*:

el alma, forma del cuerpo vivo, es una parte del cuerpo y la materia es la otra parte. Aquí la palabra “parte” se usa en un sentido especial que parece el más propio para este contexto. (Kenny 147)³

De todas maneras, más adelante volveré sobre el tema de la *parte*.

II

El examen de la cuestión 76, en su primer artículo, refuerza las ideas expuestas y nos da una explicación más metafísica y radical de la unidad de la persona. Aquí la cuestión que se plantea es si el principio intelectual se une al cuerpo como forma. Sabemos que esta es la raíz de la explicación aristotélica, a la que el Aquinate añadirá importantes elementos metafísicos que ayudarán a resolver dificultades que Aristóteles, sólo mencionó y dejó como cuestiones abiertas a saber: la posibilidad de la inmortalidad del alma y la cuestión de la unión de un principio incorruptible con uno corruptible como lo es el cuerpo.

El primer argumento que revisaré, sostiene que el alma es en la persona principio de la operación intelectual es forma del cuerpo humano. Esto se demuestra, debido a que lo primero por lo que obra un ser es su forma. Ningún ser obra, sino en cuanto que está en acto; obra por aquello que hace que esté en acto: esto es el alma. En la persona, el alma es el principio primero por el que vivimos, sentimos y entendemos primaria y radicalmente (Aristóteles, *De anima* II 2 414a 12), de ahí que podamos también derivar que la forma del hombre es una forma intelectual.

En este punto, Kenny plantea la siguiente dificultad a esta argumentación: si atendemos a la frase “*id quo primo aliquid operatur*”, nos encontramos con la expresión *id quo*, formulada en el caso ablativo latino que se usa instrumentalmente para referirse a algo hecho por un agente, por medio de una cosa diferente de sí mismo. Kenny argumenta que: “al hablar del actuar de un agente “con” o “en virtud de” su forma, se plantea un problema” (Kenny 154).

³ Uso la traducción castellana de la edición de Herder, 2000

La tensión, señala Kenny se produce porque Tomás de Aquino maneja una noción que se podría llamar abstracta de forma y otra en la que la forma actúa como agente. En un caso se usa la causa formal y en el otro la eficiente. En el primer caso, el alma es entendida como una configuración o estructura en el segundo caso como agente y responsable de las diversas actividades de la vida de la persona humana (Kenny 157). Ambas maneras de entender la forma, introducen una dificultad a la hora de explicar el alma y su relación con el cuerpo y con sus operaciones. Para Tomás de Aquino el sentido principal de la forma, designa un principio de las acciones que no es instrumental, sino que procede del interior del sujeto. El conocimiento no se hace con algo, sino en virtud que el sujeto es cognoscente. Sin embargo, argumenta Kenny, no está nítidamente trazada la distinción y no por una razón exclusivamente idiomática, sino que no queda clara la diferencia entre la relación que se establece entre un agente y un instrumento como es el caso del hombre que corta el pan con un cuchillo y la relación entre un agente y la forma en virtud de la cual, realiza la acción, como sucede en el caso del conocimiento:

No parece fácil conciliar lo que aquí se dice; a saber, que el alma es el “principio con el que primero pensamos”, y la declaración anterior sobre el conocimiento, designado como principio primero con el que el alma conoce. (Ibíd. 155)

Me parece que la cuestión planteada por Kenny se aclara con la continuación de la argumentación del artículo, pero quisiera desde ya, decir algunas cosas que llevarán a plantear el asunto de una manera diferente.

Lo primero que habría que decir es que el problema de un uso idiomático ablativo del *id quo*, no presenta mayores dificultades, porque es frecuente usar traslaticamente expresiones como éstas. Por ejemplo, cuando digo: “ejerzo la caridad por medio de una buena obra”, no estoy haciendo un uso instrumental de la caridad, sino expresando la acción, por la cual la virtud se realiza. Me parece que en Tomás de Aquino no hay tal confusión en el caso del alma, debido a que él no se refiere al alma o al entendimiento como “instrumento” con el que entendemos, sino que siempre considera el alma y el entendimiento como principio primero constitutivo de la naturaleza humana. Sólo desde esta perspectiva, comprende el papel del entendimiento y del alma en términos operacionales. Lo dice en el *corpus* del artículo:

Es manifiesto que el cuerpo vive por el alma. La vida se manifiesta según las diversas operaciones en diversos grados en los seres vivos, y el alma es lo primero por lo que estas operaciones se ejecutan. En efecto, el alma es lo primero por lo que nos alimentamos, sentimos y nos movemos localmente; así mismo es lo primero por lo que entendemos. Por lo tanto, este principio por el que primeramente entendemos, tanto si le llamamos entendimiento como alma intelectual, es forma del cuerpo. (Aquino, *Suma Teológica* c. 76 a. 1)

Este texto, nos lleva a situar la cuestión, no en el plano de una causalidad eficiente que podría entenderse de un modo instrumental, más bien, en el de la causa formal: es el ser mismo del alma el que explica sus operaciones, actuando efectivamente, como primer principio de lo que son las operaciones propias de los diversos seres vivos. Hay que advertir, también con relación al texto citado que la relación alma-cuerpo es diferente de la relación alma-conocimiento, entendido éste como operación del alma, debido a que el cuerpo posee la vida en virtud de algo diferente de él mismo, que tiene la vida como perfección propia, en cambio el alma posee el conocimiento como operación propia y no en virtud de otra cosa. No se trataría, como lo plantea Kenny, de una propiedad en virtud de la cual el ser vivo, en este caso la persona humana que actúa pudiera ser entendida instrumentalmente, sino de algo más radical: un principio por el que se *es* de un determinado modo que regula o explica el operar de ese peculiar modo de ser.

El texto recién citado, está en perfecto acuerdo con el anterior (cf. Aquino, *De anima* 2 414a 12), que es sumamente interesante al decir que el alma es aquello por lo que *primaria y radicalmente* ejercemos las operaciones propias. Es lógico que en las líneas siguientes a este texto, Aristóteles concluya que por esta razón el alma es forma específica (línea 14) (cf. Aquino, *De anima* II lec. 4 nn. 273-74). A este aspecto hace alusión como de pasada Tomás de Aquino en el comienzo del artículo cuando dice que: “el entender es propio del hombre”. Es decir, no se trata de una operación entre otras que posee el hombre, sino de aquella que lo especifica.

De esta manera, el entender se explica en conexión directa con la forma. No es producto de una causalidad eficiente, sino siempre formal por lo cual no se daría una contraposición entre una consideración abstracta y otra eficiente del alma.

III

Siguiendo con el artículo 1º, de la cuestión 76, Tomás de Aquino ratifica la afirmación inicial que el alma intelectual, es forma del cuerpo con un hecho; el entender es una acción de este hombre concreto, porque quien entiende, experimenta que es él quien entiende:

Si alguien insiste en afirmar que el alma intelectual no es forma del cuerpo, es necesario que señale cómo el entender es una acción de este hombre concreto; pues el que entiende experimenta que él es el que entiende. (Aquino, *Suma Teológica* I 76 1)

El hecho confirma el principio metafísico que aquello por lo que primero algo opera es su forma. Las operaciones son actualidades secundarias y derivadas que exigen como fundamento un acto sustancial: el alma (Guiu 175). Pero más aún en este caso, debido a que el entender es la operación propia del hombre y su operación más perfecta⁴.

Sin embargo, hay que explicar de qué manera el entender es una acción de este hombre concreto. Para esto Tomás de Aquino examina tres posibilidades de atribuir a un principio la acción: i) algo actúa o mueve, con todo su ser; ii) con parte de su ser; iii) por medio de algo accidental.

La tercera posibilidad se descarta inmediatamente, porque el entender como ya se ha expuesto es algo esencial al hombre. La posibilidad (i), que habría sido sostenida por Platón, significaría negar el cuerpo y decir que el hombre es el alma intelectual. Esta postura ya se discutió y descartó en la c. 75, a. 4. Sólo se sostiene (ii), la cual admite que el hombre entiende con parte de su ser:

Por lo tanto, hay que concluir que el entendimiento con el que Sócrates entiende, es alguna parte de Sócrates; así como que el entendimiento de algún modo está unido al cuerpo de Sócrates. (Aquino, *Suma Teológica* I 76 1)

Tomás tiene cuidado en afirmar, tanto que el entendimiento es parte del ser de Sócrates como también que esta parte está unida al cuerpo. Sin embargo, esta afirmación, no deja de plantear algunas dificultades. La

⁴ Hay que tener en cuenta que, el principio intelectual en el lenguaje tomasiano, comprende también, la voluntad.

primera que podemos señalar es que si el conocer es esencial al hombre cuesta comprender con claridad que sea llevado a cabo con parte de su ser. Ahora bien, la comprensión de la persona como formada de *partes*, dificulta la auténtica interpretación de la unidad alma-cuerpo. Si el alma (y con ella el principio intelectual) es forma, ¿cómo puede explicarse el conocer como la operación de una parte? Ciertamente, la dificultad se plantea por la característica peculiar de la persona humana, compuesta por dos principios constitutivos, pero también porque la terminología que usamos para explicar esta composición está permanentemente planteando una cierta división. Para poder responder satisfactoriamente a esta cuestión, tenemos que dar un paso más en la explicación que desarrolla el Aquinate y profundizar en la índole metafísica del alma como forma.

IV

Tomás de Aquino examina de una manera crítica dos posturas para explicar cómo se vincula el alma con el cuerpo en el hombre: la sostenida por Averroes y la atribuida por Aristóteles a Platón del alma como motor del cuerpo. No me detendré en la primera postura, sino más bien, en la consideración del alma como motor del cuerpo y en la posición aristotélica que explica que este hombre entiende porque el principio intelectual es su forma, la cual se une al cuerpo justamente como tal, pero no conoce por medio del cuerpo como si éste fuera instrumento puesto que el intelecto carece de órgano corpóreo⁵.

Antes de desarrollar estas ideas, examinemos algunas dificultades que plantea Kenny las cuales van en la misma línea de la comprensión del alma como instrumento o agente expuesta en el primer argumento. La tensión entre la noción abstracta de alma y la causalidad eficiente que se le atribuye, vuelve a aparecer al considerarse el alma como principio de vida, pero también como motor del cuerpo, aunque esta condición no sea suficiente para explicar su unión con el cuerpo. El alma es a la vez aquello que puede subsistir por sí mismo y la forma del cuerpo, haciéndolo ser un cuerpo humano (Kenny 156-57).

⁵ Aristóteles, no expresa en ningún texto, la afirmación que, es *este* hombre, el que entiende, como le hace decir Tomás de Aquino, de ahí que se hayan suscitado tantas controversias, a propósito de lo que realmente significa, el intelecto separado del *De anima* III, 5. Tomás de Aquino, siempre interpretó este texto, como una afirmación que, la operación intelectual la realiza el hombre, sin la intervención del cuerpo, rechazando la posibilidad de concluir la existencia separada del entendimiento (cf. *III De Anima*, lec. 10. nn. 734-36).

Efectivamente, la demostración que el alma es principio de vida de los vivientes y que es acto y forma de los cuerpos vivientes, es una demostración a partir de los efectos (las operaciones vitales), los cuales conducen a la causa de estos efectos: la forma del cuerpo viviente. Surge la dificultad de si al hablar de efectos, no estamos en terreno de la causa eficiente, coincidiendo ambas causalidades como sucede en la perspectiva aristotélica. El problema, está más bien, en un nivel de descripción de la operación, debido a que ésta puede considerarse, tanto en un nivel universal, como individual.

Si consideramos en el obrar humano su actividad intelectual, estamos ciertamente aludiendo a la causa formal, porque éste es el principio intrínseco en virtud del cual la persona obra, porque es también el principio por el cual es lo que es como he dicho. Pero si atendemos al acto cognoscitivo del sujeto, por ejemplo, Sócrates la causa se transforma en eficiente, porque es la operación de Sócrates, la que produce el acto cognoscitivo y en este caso es Sócrates en cuanto agente el que explica su acto de conocer.

La relación en el nivel de la causa formal, sería una relación del alma como principio vital, respecto al cuerpo y del alma con respecto a las operaciones propias del individuo. En cambio, en el nivel de la causa eficiente, la relación sería del alma, de este sujeto, a esta operación de este sujeto.

Quizás ayude a aclarar este punto, revisar la contra argumentación de Tomás de Aquino respecto a la postura que considera el alma como motor del cuerpo. Tal como dice Kenny, lo que Aquino rechaza es que desde esta posición sea posible sostener que el alma se una al cuerpo y se salve la unión sustancial de ambos principios, unión que finalmente, explica que es *este* hombre, el que entiende y algo más importante que en definitiva, podamos hablar del hombre como un ser sustancial, debido a que “cada cosa es ser en tanto que es una” (*Suma Teológica*, 1 76 1 *corpus*).

Atendamos al segundo argumento que presenta, el cual es resumidamente el siguiente: si se considera el alma o el entendimiento como motor del cuerpo, el entendimiento resultaría un principio extraño a la esencia del hombre, porque el motor actúa por así decirlo desde fuera del móvil o al menos, no es necesario que pertenezca a la esencia del móvil. Esta posibilidad es rechazada por dos motivos: i) porque entender, es una operación esencial a la persona, como hemos visto; ii) el entender, es

una operación inmanente y tales acciones no se explican por una causa agente, sino por la forma del que actúa.

La postura aristotélica resuelve estas dificultades, porque sostiene la unión del entendimiento al cuerpo como forma y así explica que la operación de entender sea la operación esencial del hombre y que se atribuya también a *este* hombre. Como dice Tomás de Aquino en el *corpus* del artículo, la naturaleza de cualquier cosa queda manifiesta por su operación y la operación propia del hombre en cuanto hombre es la de entender, operación que realiza en virtud de su forma propia y sustancial, es decir, aquella forma que le confiere unidad en el ser; finalmente, explica también, que se trata en este caso de una operación inmanente, debido a que las operaciones que proceden de la forma son inmanentes al sujeto.

Por esto me parece, a diferencia de Kenny y como ya expliqué en el primer argumento que toda esta explicación tomasiana, considera fundamentalmente la noción de forma sustancial y que cuando maneja la causalidad eficiente, por ejemplo, para explicar las operaciones del sujeto, supedita esta causalidad a la formal. En este sentido, digo que considera *fundamentalmente* la noción de forma.

V

Sin embargo, aún resta por resolver una dificultad que dejamos pendiente y es explicar de qué manera si esta unión es total -sustancial- el hombre entiende con *parte* de su ser, es decir, con el entendimiento que no es todo el ser. Examinemos este punto.

En la última parte del *corpus* del artículo, Tomás de Aquino expone lo que es el elemento innovador con respecto a la postura aristotélica y también la explicación más radical de la unidad alma-cuerpo sin la cual todas las otras explicaciones, resultan en cierto modo insuficientes o podrían presentar un flanco de posibles objeciones. Me refiero a la doctrina tomasiana del alma como forma emergente de la materia. Al hablar de emergencia, no me refiero a las teorías emergentistas del alma que la consideran como un epifenómeno de la materia, sino exclusivamente al hecho de que el alma tiene una actividad que supera la actualización de la materia. Es decir, que el alma comunica el ser al cuerpo su realidad no se agota en la actualización del cuerpo, sino que tiene una perfección que le permite actuar, también por sí misma.

Tomás de Aquino, reconoce la desconfianza de los padres de la Iglesia, concretamente Gregorio de Nisa (cf. Tomás de Aquino, *Las criaturas espirituales* c. a. 2), respecto a la concepción aristotélica que parecía comprometer la espiritualidad del alma. La razón principal, es que si se define el alma como acto primero del cuerpo, ésta tendría que ser mortal, porque nada incorruptible puede ser forma de un cuerpo corruptible, debido a que a la forma le compete de suyo ser en una materia y nada puede ser sin aquello en lo que inhiere propiamente. De ahí que la forma del cuerpo se destruiría una vez destruido el cuerpo o bien pasaría a ser otro cuerpo (cf. Tomás de Aquino, *De la unidad del entendimiento* c. 1 nn. 15-6).

Metafísicamente lo que late en el fondo de esta polémica es que aquello que existe en la materia, es material, porque no es distinto el ser de la forma y el ser de la materia. Esta es una objeción seria desde el punto de vista metafísico a la unión sustancial de alma y cuerpo. Es esta dificultad la que permite comprender la solución de Averroes, debido a que si el principio intelectual está realmente separado del cuerpo se salva su espiritualidad (cf. Tomás de Aquino, *Las criaturas espirituales* c. a. 2).

La solución a esta dificultad, sólo se puede comprender desde la vía del ser y de la emergencia de la forma respecto a la materia como sostiene Tomás en el artículo que estamos analizando. Me parece importante citar el texto en su integridad:

Sin embargo, hay que tener presente que una forma, cuanto más alta es su categoría, tanto más domina la materia corporal y menos inmersa está en ella, y tanto más la impulsa por su operación y su capacidad. Así observamos que la forma de un cuerpo compuesto tiene alguna operación que no es causada por las cualidades fundamentales. Cuanto mayor es la categoría de las formas, tanto más supera su poder al de la materia elemental; y, de este modo, el alma vegetativa supera la forma de un metal; lo mismo hace el alma sensitiva respecto de la vegetativa. Pero de todas las formas, la de más categoría es el alma humana. Por eso, su poder sobrepasa de tal manera al de la materia corporal, que tiene una capacidad y una operación en la que de ninguna manera participa la materia corporal. Esta facultad es llamada entendimiento. (Tomás de Aquino, *Suma Teológica* c. 76 a. 1)

Las objeciones que se han planteado en el artículo mencionado, tienen en común señalar la dificultad de la unión del alma como forma del cuerpo, porque de ahí se seguiría la negación del carácter separado del entendimiento y su capacidad universal de conocer. Si consideramos también las respuestas a estas objeciones, podemos completar esta explicación, más allá de la categorización superior del alma humana, en relación a las almas de los otros seres vivos e incluso, a la forma de otros seres materiales.

Tomás de Aquino añade en la respuesta a la objeción quinta, una idea que es crucial, a saber que es el alma la que comunica su ser a la materia, siendo el ser común al compuesto, es decir, el ser de la persona abarca cuerpo y alma. Junto con la emergencia del ser del alma se pueden explicar, tanto la perpetuidad del alma una vez corrompido el cuerpo, como también la capacidad del entendimiento de realizar operaciones propias.

Esta explicación, sirve también para solucionar las dificultades que se presentaban al considerar que el entendimiento es separado y que lleva a decir que el hombre entiende con una *parte* de su ser y no con el cuerpo. Se entiende por qué usa el término *parte*. Decíamos en su momento que esta forma de expresarse del Aquinate, ponía de cierta manera en cuestión la unidad alma-cuerpo. Pero ahora, estamos mejor preparados para enfrentar esta dificultad. La separación del entendimiento con respecto al cuerpo, no es en su calidad de forma sustancial -en el aspecto en el cual el entendimiento es asimilado en el lenguaje tomasiano al alma-, sino más bien en cuanto capaz de realizar operaciones propias, es decir, en cuanto facultad del alma, porque el alma que es forma sustancial del hombre se une indisolublemente al cuerpo formando una unidad.

VI

Este aspecto de la unión sustancial, está suficientemente explicado en el *corpus* y reforzado en las respuestas a todas las objeciones. Vemos que aunque se le adjudica al entendimiento una operación propia y separada del cuerpo, siempre se acompaña esta afirmación de la comprensión del alma intelectual como forma del cuerpo, del cual no está separada:

En efecto, está separado por su potencia intelectual, porque el poder intelectual no es facultad de ningún órgano corpóreo, tal como la facultad de ver es acto de los ojos; pues entender es un acto que no puede ejercerse mediante algún órgano corporal, como se realiza el acto de ver. Sin embargo, está en la materia por cuanto que la propia alma, a quien pertenece tal facultad, es forma del cuerpo y término de la generación humana. (Tomás de Aquino, *Suma Teológica* c. 76 a. 1 ad 1)

El entendimiento, no es considerado acto del cuerpo en su calidad de facultad, sino como el principio formal que constituye a la persona humana en cuanto tal; la facultad cognoscitiva y el principio formal (el alma) están estrechamente vinculados, pero se los puede diferenciar según se atiende a las operaciones o al ser propio de la persona humana (Ibíd. ad. 2 3 4).

En la respuesta a la objeción cuarta, la cuestión es tratada de una manera completa, debido a que se afirma, tanto la emergencia del alma en relación al cuerpo, como también su calidad de forma del cuerpo.

Como una conclusión del modo de existir que tiene el alma humana y de su vinculación con el cuerpo, Tomás de Aquino establece un orden de perfección entre las diferentes almas que informan a los seres animados. Esta perfección se funda en el mayor dominio que la forma tenga con respecto a la materia, su mayor *emergencia* con respecto a ella. Así, concluye que la de más categoría (*nobilior*) es el alma humana.

Kenny objeta esta afirmación y sostiene que apelar a una jerarquía o nobleza entre las formas es una forma de argumentación difícil de sostener actualmente en alguna escuela: “el problema de las formas sin materia no es cuestión de valor, sino de lógica” (Kenny 159).

Por tal motivo, me parece que es difícil comprender esta afirmación de Tomás desde una perspectiva que no sea la ontológica, porque la mayor nobleza del alma está sostenida por criterios de prioridad ontológica. La mencionada superioridad del alma humana se manifiesta en unas operaciones que pueden ser evaluadas como más nobles respecto de otras, porque son capaces de una mayor o más perfecta interacción con el mundo. Así, por ejemplo, la vida sensitiva es superior al estado mineral, no por una mera valoración subjetiva, sino por una forma de operar del

ser vivo sensitivo que lo sitúa en un nivel de interacción con el mundo, objetivamente superior al del mineral. Opino que una característica central del pensamiento de Tomás de Aquino es apoyar los juicios de valor en perfecciones entitativas, las que a su vez se manifiestan en el comportamiento de los entes, es decir, son manifiestas en las operaciones correspondientes a la forma que el ente posee. Desde esta perspectiva se puede decir que toda su filosofía es una explicitación de la perfección del ser como la más excelente y la fundante de toda otra perfección.

Conclusión

Aunque Kenny tiene un acercamiento al pensamiento de Tomás de Aquino que podríamos calificar de abierto y benevolente no logra integrar en los análisis de los textos tomasianos los elementos metafísicos que son claves para una correcta comprensión de su filosofía. Su posición es de alguna manera ambigua porque junto con valorar los aportes de Tomás a la discusión contemporánea, objeta conclusiones, que son centrales en la doctrina tomasiana del alma, debido a que no considera el alcance metafísico de ellas.

No es el caso desarrollar aquí los fundamentos ontológicos de la comprensión del alma humana que sustenta Tomás de Aquino. Brevemente, quiero destacar dos ideas que a mi modo de entender son centrales para la comprensión correcta de su postura: i) la consideración principal y fundamentalmente formal del alma y ii) la incorporación de la doctrina del ser a esta explicación. Así, puede afirmar que en el caso de la persona humana es el alma la que posee un ser propio, el cual al comunicarse al cuerpo como acto permite entender, tanto la unión sustancial con éste, como también la emergencia de ciertas operaciones que se realizan sin el cuerpo.

Ciertamente, quedan abiertas muchas otras cuestiones de sumo interés para el filósofo contemporáneo, como por ejemplo, el real alcance y sentido de esta operación separada del cuerpo que realiza el entendimiento. Problemas como éste son ampliamente debatidos en el ámbito de la filosofía de la mente, sin embargo, para llevar a cabo un diálogo provechoso con estas corrientes del pensamiento actual, es necesario tener como punto de partida, claridad respecto al pensamiento del Aquinate, quien puede aportar muchos y valiosos elementos en esta comunicación de enfoques diversos.

REFERENCIAS BIBLIOGRÁFICAS

Aquino, Tomás. *Comentario al "Libro del Alma" de Aristóteles*. Trad. María Donadío Maggi de Gandolfi. Buenos Aires: Fundación Arché, 1979. Impreso.

---. *Suma Teológica*. Madrid: BAC, 1988. Impreso.

---. "El ser y la esencia". *Opúsculos y cuestiones selectas I*. Madrid: BAC, 2001. Impreso.

---. "Las criaturas espirituales". *Opúsculos y cuestiones selectas I*. Madrid: BAC, 2001. Impreso.

---. "De la unidad del entendimiento". *Opúsculos y cuestiones selectas I*. Madrid: BAC, 2001. Impreso.

---. *Suma contra gentiles*. Madrid: BAC, 2007. Impreso.

Aristóteles. *De anima*. Trad. Tomás Calvo Martínez. Madrid: Gredos, 1988. Impreso.

---. *Metafísica*. Trad. Valentín García Yebra. Madrid: Gredos, 1990. Impreso.

Averroes. *Aristotelis opera cum Averrois commentariis*. New York: Kessinger Publishing LLC., 2010. Print.

Cayetano. *Comentarios al Del ente y la esencia*. Trad. David García Bacca. Caracas: Universidad Central de Venezuela, Ediciones de la Biblioteca, 1974. Impreso.

Guiu, Ignacio. *Sobre el alma humana*. Barcelona: PPU, 1992. Impreso.

Kenny, Anthony. *Tomás de Aquino y la mente*. Trad. J. M. López de Castro. Barcelona: Herder, 2000. Impreso.

mph

DISCUSIONES

Discussions

DO WE NEED TO REFORM THE OLD T-SCHEME?

¿NECESITAMOS REFORMAR EL VIEJO ESQUEMA T?

JAN WOLEŃSKI

University of Management and Information, Rzeszów, Finlandia. wol@if.edu.pl

RECIBIDO EL 30 DE ABRIL DE 2012 Y APROBADO EL 7 DE MAYO DE 2012

RESUMEN ABSTRACT

Este artículo comenta de manera crítica el análisis de Stephen Read del esquema T ofrecido en su artículo: "The Liar and the new T-scheme". Sostengo que Read malinterpreta los planteamientos de Tarski en algunos puntos y que el esquema A que introduce Read en su artículo, es o bien reducible al viejo esquema T (tarskiano) o bien plantea una serie de dudas esenciales.

This paper critically comments Stephen Read's analysis of the T-scheme offered in his paper: "The Liar and the new T-scheme". I argue that Read misinterprets Tarski in some points and that the A-scheme introduced in Read's paper is either reducible to the old (Tarskian) T-scheme or raises essential doubts.

PALABRAS CLAVE KEY WORDS

Falsedad, inconsistencia, intencionalidad, probabilidad, Read, Tarski, tautología, el mentiroso, verdad, oraciones T.

Falsehood, inconsistency, intensionality, provability, Read, Tarski, tautology, the Liar, truth, T-sentences.

Read 2010 almost verbatim reproduces Read 2008. I was invited to make comments about the later paper (Woleński). In what follows I partly repeat my earlier critical comments about Read's interpretation of the T-scheme, but I also add new remarks and improve old ones; some of additions are inspired by other contributors participating in Rahman, Tulenheimo, Genot (2008), including Stehpen Read's replies in this volume (2008) as well as the previous discussion in *Discusiones Filosóficas* (Miller, Read, Sandu).

At first, let me remind Leśniewski's-Tarski's diagnosis of the Liar paradox. They pointed out (Tarski 1944) that the derivation of the paradox uses: I) self-referential sentences asserting semantic properties; II) the T-scheme, and III) classical, that is, bivalent logic. Hence, we can conceive three strategies in order to solve the Liar: i) to exclude self-referentiality; ii) to reject or modify the T-scheme; iii) to change logic. It would be mistaken to maintain that there is a solution free of costs or some artificiality. It concerns Tarski who choose i), Kripke (and many other logicians) who opted for ii) (Sandu makes several remarks about this way out), and Read who tries to reform the T-scheme. More precisely, Read argues that the old (Tarskian) T-scheme is inaccurate and proposes its modification. I would like to show that Read's reading of Tarski is incorrect and that his (Read's) leads to some difficulties.

Read says that, according to Tarski, every instance of the T-scheme, that is the formula

(T) x is true if and only if p ,

is true; Read even says that his understanding of (T) is "an unquestioned orthodoxy". However, Tarski focused on the provability of T-sentences from his truth-definition, but not on their truth. In Read's reply to my criticism (Read, *The Truth-Schema* 218), he agrees with my standpoint, but he adds "But he [Tarski] clearly imposed that requirement because he thought those instances were true". I should note that Read uses the proper formulation in another place of his main paper 2010 (*The Liar and* 122). He probably thinks that both formulations are equivalent. However, they are not, unless we assume that the metatheory of truth-theory is ω -complete. Anyway, Read's statement about "an unquestioned orthodoxy" is certainly incorrect. For example, Miller and Sandu in their contributions published in *Discusiones Filosóficas* formulate (T) via provability, not truth.

I see no place in Tarski's writings which could justify the view that he "clearly imposed, etc." The observation that provable instances of (T) are true is trivial and has no relevance for Tarski's proposal of how to solve the Liar paradox. On the other hand, the provability of T-sentences matters very much and the role of this fact has the best illustration in the problem of the definability of the T-predicate. Combining the fixed-point theorem and the Tarski undefinability theorem, leads to the unprovability of some instances of (T). This fact concerns the Liar sentence, independently of its formulation as 'this sentence is false' or 'this sentence is not true', both recorded via arithmetization.

The situation is clear in formal arithmetic of natural numbers, but the results that hold for arithmetic cannot be directly applied to natural language. However, *pace* Tarski, excluding the Liar sentence from the stock of permitted formulas of natural language plays a quite similar role to showing that some formalized instances of (T) are not provable. In fact, this is an analogical move as the exclusion of division by 0 in arithmetic.

Although the formula $m/0 = n$ is grammatically correct and perfectly understandable, it must be rejected as producing inconsistency. As far as I know Tarski never said that the liar sentence is nonsensical, meaningless, etc. He only recommended that so-called closed languages, that is, languages containing semantic concept self-referentially used, should be avoided. This restriction is by no means that counterexamples to (T) are excluded by fiat as Read suggests (127). On the contrary, they are avoided by a subtle and elegant reasoning.

That T-sentences are assumed to be provable is important for their status. Read says (*The Liar and* 125) that according to Tarski (T) is "a merely a material equivalence". Consequently, the instances of (T) have the same status. Although Tarski himself was not quite explicit about this issue, it is rather obvious that provable theorems of the form $A \leftrightarrow B$ are something more than material equivalences. In particular, one cannot replace A or B by their material equivalences. For instance, the equivalence: 'the sentence 'Stephen Read wrote a paper on T-scheme' is true if and only if Stephen Read wrote a paper about T-scheme' cannot be replaced by the equivalence: 'Stephen Read wrote a paper about T-scheme if and only if Jan Woleński commented Read's paper', although both equivalences are true, and, thereby, materially equivalent, and, moreover, both consist of true constituents.

I do not suggest that Read considers such a replacement as possible or justified. Yet I claim that the difference between merely material equivalences and provable material equivalences is important, particularly for a proper interpretation of Tarski's truth-theory.

Read entirely neglects some properties of languages for which the semantic concept of truth is defined. Firstly, such a language \mathcal{L} is formalized or at least it has specified structure; this latter concept was introduced in Tarski 1944. Without entering into details, \mathcal{L} must be well-described as a syntactic object. In particular, we should know what belongs to the vocabulary of \mathcal{L} and how the class of its sentences is defined. Secondly and more importantly, \mathcal{L} is an interpreted language. Let me quote the following words of Tarski (1933):

It remains perhaps to add that we are not interested here in 'formal' languages and sciences in one special sense of the word 'formal', namely sciences to the signs and expressions of which no meaning is attached. For such sciences the problem here discussed has no relevance, it is not even meaningful. We shall always ascribe quite concrete and, for us, intelligible meanings to the signs which occur in the language we shall to consider. The expressions which we call sentences still remain sentences after the sign which occur in them have been translated into colloquial language. The sentences which are distinguished as axioms seem to us to be materially true, and in choosing rules of inferences we are always guided by the principle that when such rules are applied to true sentences the sentences obtained by their use should also be true. (166-67)

Tarski did not explain what meaning was for him. However, it is clear that meanings of words, ordinary or artificial, for instance, the convention that 'black' expresses the property of being white, dictate semantic interpretations. Speaking more precisely and using the model-theoretic terminology, meanings generate so-called interpretation functions which ascribe denotations of individual terms and predicates in given models.

Although Tarski was skeptical as far as the matter concerns consistent semantic constructions for natural language, the difference between it and interpreted formalized (or with a specified structure) \mathcal{L} is less radical than it is frequently assumed by interpreters and critics of Tarski.

Taking facts about \mathcal{L} into account one can easily demonstrate that all of the counterexamples given by Read fail. This concerns the sentences:

- (1) 'I am tired' is true if and only if I am tired;
- (2) 'That book was stolen' is true if and only if that book was
stolen;
- (3) 'Any man is mortal' is true if and only if any man is mortal.

The argument is that the words 'I', 'that' and 'any' can have different meanings on the left and right sides of the respective equivalences. However, this cannot happen by definition in the case of properly interpreted languages. In particular, such languages do not require a special principle of uniformity (Dutilh Novaes) blocking ambiguities and other defects of expressions, because they are automatically excluded by technology of the semantic interpretation.

The valuation is a function and 'I', 'that' and 'any' have to have the same interpretations in all their occurrences in (1)-(3). This is particularly clear for indexicals, like 'I' and 'that'. An ostensive specification of references (for example, supplementing the act of using 'I' or 'that' by a gesture) fixes the ascribed objects.

On the other hand, if the references of such words are not made precise, they function as variables and the problem of truth (falsehood) of (1)-(3) does not arise at all. I do not argue that ambiguities do not occur in ordinary language, but only point out that the semantic definition of truth assumes that \mathcal{L} for which it works is formalized (or has a specified structure) and its interpretation is fixed, even when \mathcal{L} is selected as a suitable (not closed!) part of colloquial parlance.

Read proposes to replace (T) by

$$(A) \quad Tx \Leftrightarrow \forall p(x : p \rightarrow p).$$

This scheme codes an intuition expressed by

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

Since what a sentence says is covered by all its implications, (A) formalizes (S), provided that the symbol \Leftrightarrow denotes strict equivalence.

Read says (*The Liar and* 125) that “(A) is a logical equivalence” (expressed by the symbol \Leftrightarrow) contrary to (T) as a material equivalence.

However, Read’s qualification of (A) as a logical equivalence is vague. If we say that A is a logical sentence, we can mean various things. Firstly, (a) A is logical if and only if it is coded by signs belonging to the language of logic, but, secondly, (b) A is logical if and only if it is a logical theorem.

Now, (T) and (A), are logical equivalences in the same sense, if the first understanding is assumed, but they both are not logical theorems.

Thus, neither (a) nor (b) qualify (A) as a logical equivalence, contrary to (T). Perhaps Read intends to say that (A) is a logical equivalence, because strict connectives are logically stronger than material ones. However, due to my previous remarks about the status of the instances of (T), provable equivalences are “something more” than material equivalences. In fact, the difference between the formulas $X \vdash A \leftrightarrow B$ and $A \leftrightarrow B$ seems secondary, although, at least in my view, the former is much clearer than the latter.

Read points out that (A) is intensional for ‘says that’, but he does not see any problem with it. In particular, he seems to think that everything is solved by the closure of $x : p$ by “allowing substitution only of logical equivalents” (Read, *The Liar and* 124).

Unfortunately, the matter is not so simple, because ‘says that’ is strongly intensional. Consider two equivalent formulations of a mathematical axiom, let say, the parallel postulate. Denote them by F and F' , respectively. They are logically equivalent. Assume that we have a person O who does not know that F and F' are provably equivalent. Thus, the formulas $x : F$ and $x : F'$ do not say the same for O , although Read claims that they do.

Otherwise speaking, ‘says that’ forms contexts which are, contrary to Read, not closed under replacement by logical equivalents. Defining ‘says that’ as satisfying such a kind of closure is, in my opinion, at odds with the ordinary meaning of this operator.

There are other implausible consequences of (A). Consider (I do not assume that Peano arithmetic is first-order)

$$(4) \quad \mathbf{T}(\text{'Peano axioms'}) \Leftrightarrow \forall p (\text{'Peano axioms': } p \rightarrow p).$$

Peano axioms are true in standard and non-standard model of arithmetic. Let \mathbf{N} be the standard model and \mathbf{N}' a nonstandard one. Take the sentence (*) 'all natural numbers have finitely many predecessors', which is true in \mathbf{N} , but false in \mathbf{N}' . Intuitively speaking, Peano axioms say (*) in \mathbf{N} , but its negation in \mathbf{N}' .

Thus, 'says that' requires a relativisation to a model, but **(T)** without it has false instances. This also shows that defining the context $x : p$ by the consequences of p may be insufficient in some cases, because a reference to models is required.

There is also a problem with **(A)** as applied to falsehoods (this is also pointed out by Miller and Sandu). The definition of **F** ('is false') corresponding to **(A')** can be recorded as

$$(\mathbf{A}'') \mathbf{F}x \Leftrightarrow \exists p ((x : p) \wedge \neg p).$$

As an example we have

$$(5) \mathbf{F} (\text{'Warsaw is the capital of France'}) \Leftrightarrow p (\text{'Warsaw is the capital of France': } p) \wedge \neg p).$$

However, (5) has its instantiation in

$$(6) \quad \mathbf{F} (\text{'Warsaw is the capital of France'}) \Leftrightarrow ((\text{'Warsaw is the capital of France'}): \text{the greatest Polish city is the capital of France}) \wedge \text{the greatest Polish city not is the capital of France}.$$

Now assume that someone knows that a person O knows that the sentence 'the greatest Polish city is the capital of France' is false and that the right side of (6) is a correct (true) instantiation of the right side of (5).

Hence, he knows that the sentence 'Warsaw is the capital of France' is also false; we use here the principle 'if an instantiation of A is false, A is false too. However, to justify that, one must assume that 'Warsaw' and 'the greatest Polish city' are co-denotative (I neglect that the latter is a description). This consideration shows that we do not need to worry whether the sentences 'Warsaw is the capital of France' and 'the greatest Polish city is the capital of France' say the same or, eventually, in which

circumstances they cover the same content, because it is sufficient to know the values of nominal expressions.

Thus, even if we agree that that logically equivalent sentences say the same thing, this observation does not close the issue, because it can happen, as in the case of (5) and (6) that sentences are equivalent in theories or some language systems modulo denotative conventions, although they are not equivalent on purely logical grounds. Thus the interpretation of a language has a crucial importance for establishing what sentences say and when they are true or false.

Further, the negation of $(x : p)$ can be interpreted either as $\neg(x : p)$ or as $(x : \neg p)$ (the latter is stronger than the former). It matters in the case of negative sentences because

(7) \mathbf{T} ('Warsaw is not the capital of France') $\Leftrightarrow \forall p (\neg p \rightarrow \text{'Warsaw is not the capital of Poland': } \neg p)$ looks more plausible than

(8) \mathbf{T} ('Warsaw is not the capital of Poland') $\Leftrightarrow \forall p (\neg p \rightarrow (\text{it is not the case that 'Warsaw is not the capital of Poland': } \neg p))$.

Perhaps the most important critical observation concerns metalogical properties of \mathbf{T} . The law of the excluded middle can be stated as

(9) $\forall p (x : p \rightarrow p) \vee \exists p ((x : p) \wedge \neg p)$,

which is an instance of

(10) $\mathbf{TA} \vee \neg\mathbf{TA}$.

However, the relation of \mathbf{TA} and $\mathbf{T}\neg A$ is not clear under Read's definition. In general we have

(11) $\mathbf{T}\neg A \rightarrow \neg\mathbf{TA}$,

and in the classical bivalent (two-valued) case the equivalence

(12) $\mathbf{T}\neg A \leftrightarrow \neg\mathbf{TA} \leftrightarrow \mathbf{FA}$

holds. Now $T\neg A$ has its interpretation (according to Read's definition) in

$$(13) \quad \forall p (x : \neg p \rightarrow \neg p).$$

However, (13) does not implies

$$(14) \quad \exists p ((x : p) \wedge \neg p),$$

although we have (or should have)

$$(15) \quad (x : \neg p) \rightarrow \neg(x : p).$$

I do not claim that (12) is indispensable, but only note that important relations are unclear under Read's proposals. In particular, the functor of negation commutes with truth (in its semantic understanding), but not with 'says that'.

Returning to 'x says that p', a full analysis of this phrase seems to be much more complicated than Reads maintains. The general problem is that sentences can say quite different things for different persons or even for the same person depending of various pragmatic factors. For instance, we should distinguish direct (explicit) and indirect (implicit saying that. Every sentence entails infinitely many logical consequences. Even if we assume that the direct linguistic content of a sentence A is definable (it is a very optimistic presumption), the implicit content is much vaguer and the amount of its grasping, always partial, cannot be accounted in advanced. Reads seems to consider the intensionality of (A) as its advantage, but it is a very dubious view.

I am inclined to think that the intensionality of (A) prevents a satisfactory definition of saying wholly as things are. This is very important, because the real virtue of the scheme introduced by Read essentially depends on such a definition.

(T) essentially differ from (A), because the former it is purely extensional. In spite of this difference, I will argue that both schemes record almost the same intuitions. It is interesting that Tarski's starting point was similar to that proposed by Read. The initial intuition of the semantic definition of truth was presented by the following formulat of Tarski (1933) (155); Tarski followed Tadeusz Kotarbiński:

- (*) A true sentence is one which says that the states of affairs are so and so, and the state of affairs indeed is so and so.

The original Polish formulation is much closer to (A), because it runs (in English translation): “a true sentence is a sentence which expresses that things are so and so and things indeed are so and so”. Tarski explicitly formalized (*) by T, presumably also for eliminating ‘expresses’ as an intensional factor. That \mathcal{L} is interpreted and that its expressions have “intelligible meanings” illuminate the issue at stake.

(T) can be supplemented by assertions of the type ‘ x says that p ’, ‘ x means that p ’, ‘ x asserts that p ’, etc. All are external with respect to instance of (T) and serve to fix, define, explain, etc. the valuation function connecting \mathcal{L} -expressions with their denotations. However, ‘ $x : p$ ’, although internally embedded into (A) plays exactly the same role, provided that semantics a la Tarski is associated with this scheme.

However, Read explicitly says (*The Liar and* 125) that this is just the case. In my opinion, (T) has similar intuitive advantages as (A), including the correspondence platitude, that is, the ability to express the basic content of *adequatio rei et intellectus*.

On the other hand, the Tarski scheme avoids all problems caused by the internal intensionality of (A), in particular, the treatment of ‘is false’ (remember Russell’s requirement that any satisfactory theory of truth should be also the theory of falsehood). In order to explain this point let me refer to Miller 2010 (223). He attributes to Tarski the following scheme (I use the notation of the present paper):

$$(T') \mathbf{T}x \leftrightarrow (x \in \mathcal{L}) \wedge p.$$

The first conjunct in the right side of (T') indicates that the sentence named by x belongs to \mathcal{L} .

Miller define ‘is false’ by

$$(F') \mathbf{F}x \leftrightarrow (x \in \mathcal{L}) \wedge \neg p.$$

Since $\mathbf{T}x$ and $\mathbf{F}x$ are mutual negations in classical logic, the same relation must hold between the right sides of (T') and (F').

However, $(x \in \mathcal{L}) \wedge p$ and $(x \in \mathcal{L}) \wedge \neg p$ do not negate each other. The negation of the former formula yields the sentence $\neg(x \in \mathcal{L}) \vee \neg p$. Assume that this disjunction is true in favor of $\neg(x \in \mathcal{L})$. Now, since x does not belong to \mathcal{L} , it is not a sentence.

Consequently, it is not true and not false, because only sentences of \mathcal{L} can be true or false. Clearly, $\mathbf{F}x$ and $\neg\mathbf{T}x$ are not equivalent under (\mathbf{T}') and (\mathbf{F}') . This example additionally shows difficulties when truth is defined by conjunctions of conditions. It is much better to treat clauses, like $x \in \mathcal{L}$ as external with respect to (\mathbf{T}) .

Read claims that (\mathbf{A}) is better, because it correctly solves the Liar paradox. According to Read, (\mathbf{T}) is false, because it has a false instance (the letter l refers to the Liar sentence):

(#) l is true iff and only if l is not true.

On the other hand, (\mathbf{A}) produces a correct truth-condition of l in the form:

(##) $\mathbf{T}l \leftrightarrow \neg\mathbf{T}l \wedge \mathbf{T}l$.

However, this success is entirely apparent. The formula $(\#)$ does not give a complete analysis of the Liar. First of all, $(\#)$, as inconsistent, is not provable in the semantic theory of truth. Since $(\#)$ is true for Read, the formula $\mathbf{T}l$ is either false or inconsistent. However, both cases must be excluded. A simple transformation of $(\#)$ gives:

(####) $\neg\mathbf{T}l \leftrightarrow \mathbf{T}l \vee \neg\mathbf{T}l$,

which is not satisfactory, because it says that a tautology is not true, that is, inconsistent. In fact, adding $\neg\mathbf{T}l$ (this is an unprovable formula!) to the stock of theorems of propositional logic immediately abolishes the Post (absolute) consistency of this logic. The actual situation of l is displayed by the formula

(####) $\mathbf{T}l \leftrightarrow \mathbf{F}l \leftrightarrow \neg\mathbf{T}l \wedge \mathbf{T}l$.

Roughly speaking, every semantic assumption about l , that is, concerning its truth-value modulo the bivalence, entails a contradiction. It clearly shows, firstly, why some instances of (\mathbf{T}) are excluded as unprovable and, secondly, that $(\#)$ does not formulate any truth-condition. This

concurr with Sandu's remark that (Sandu 289), it is not clear what the Liar sentence says, although I am inclined to a more radical conclusion, namely that *l* says nothing.

To sum up, if (A) is reduced to (T), the former solves the Liar in the same way as the latter does, but if it is not reduced, the issue is still open, because the proponent of (A) must decide whether this schema has true, but unprovable instances.

Finally, I would like to note that the problem of a philosophical significance of the T-scheme is still open. Does it code the correspondence intuition or not? Should we modify (T) in order to express the correspondence platitude? I do not address to these and similar questions in this paper. I hope to give an account of the philosophical significance, if any, of Tarski's theory of truth in Woleński (in preparation).

BIBLIOGRAPHICAL REFERENCES

Dutillh Novaes, Catarina. "Tarski's hidden theory of meaning: Sentences say exactly". *Unity, truth and the Liar. The modern relevance of medieval solutions to the Liar Paradox*. Shahid Rahman, Tero Tulenheimo, Emmanuel Genot. Dordrecht: Springer, 2008. 41-64. Print.

Miller, David. "Verdades, antiguas y modernas". *Discusiones Filosóficas*. Ene.-Jun. 2011: 207-227. Impreso.

Rahman, Shahid., Tero Tulenheimo and Emmanuel Genot. Eds. *Unity, truth and the Liar. The modern relevance of medieval solutions to the Liar Paradox*. Dordrecht: Springer, 2008. Print.

Read, Stephen. "The Truth-Schema and the Liar". *Unity, truth and the Liar. The modern relevance of medieval solutions to the Liar Paradox*. Shahid Rahman, Tero Tulenheimo, Emmanuel Genot. Dordrecht: Springer, 2008. 3-18. Print.

---. "Further thoughts on Tarski's T-scheme and the Liar". *Unity, truth and the Liar. The modern relevance of medieval solutions to the Liar Paradox*. Shahid Rahman, Tero Tulenheimo, Emmanuel Genot. Dordrecht: Springer, 2008. 205-228. Print.

---. "The Liar and the new truth-Schema". *Discusiones Filosóficas*. Jul.-Dic. 2010: 119-137. Print.

---. "Miller, Bradwardine and truth". *Discusiones Filosóficas*. Ene.-Jul. 2011: 229-235. Print.

Sandu, Gabriel. "Read on the Liar". *Discusiones Filosóficas*. Jul.-Dic. 2011: 277-290. Print.

Tarski, Alfred. "The concept of truth in formalized languages". *Logic, semantics, Metamathematics*. Oxford: Oxford University Press, 1933. Print.

---. "The semantic concept of truth and the foundations of semantics". *Philosophy and Phenomenological Research*. 1944: 341-375. Print.

Woleński, Jan. "Read about T-scheme". *Unity, truth and the Liar. The modern relevance of medieval solutions to the Liar Paradox*. Shahid Rahman, Tero Tulenheimo, Emmanuel Genot. Dordrecht: Springer, 2008. 199-203. Print.

---. *Semantics and truth*. Forthcoming.

ph

COLABORADORES

COLLABORATORS

ROM HARRÉ

Matemático y físico neozelandés. Realizó estudios en filosofía y antropología, fue alumno de J. L. Austin en Oxford. Actualmente, es docente del Departamento de Psicología de la Universidad de Georgetown y Director del Centro de Filosofía Natural y Ciencias Sociales de la London School of Economics. Harré, ha escrito sobre una amplia variedad de temas, que incluyen: filosofía de las matemáticas, filosofía de la ciencia, ontología, psicología, psicología social y sociología de la filosofía. Sus publicaciones más importantes son: *The social construction of emotions* (1986), *Realism rescued: how scientific progress is possible* (1994), *Cognitive science: A philosophical introduction* (2001) y *Key thinkers in psychology* (2006).

JOHN PEACOCK

Físico de la Universidad de Glasgow, Ingeniero de la Universidad de Strathclyde. Actualmente, pertenece al Departamento de Física de la Universidad de Amberes, donde se encuentra realizando sus estudios de doctorado en el laboratorio de física biomédica. La mayor parte de su trabajo, gira en torno a la comprensión del sistema auditivo en mamíferos. Ha publicado una gran variedad de artículos en revistas especializadas a nivel mundial.

MARK A. NOTTURNO

Doctor en Filosofía de la Universidad de Columbia. Notturmo, ha ocupado cargos docentes y de investigación, en diversas universidades y fundaciones; de 1994 a 1999, fue Director de "Soros Ianus Foundations". Actualmente, es Director de "Interactivity Foundation", fundación que se dedica a generar un debate filosófico, sobre cada una de las implementaciones de políticas públicas, en los Estados Unidos. Sus

publicaciones más importantes son: *Objectivity, rationality and the third realm* (1985), *On Popper, science and the Open society, and objectivity* (2002). Además, fue el editor junto a Popper, de: *Knowledge and the Body-Mind problem* (1994) y *Popper's the myth of the framework* (1996).

TONI RØNNOW-RASMUSSEN

Doctor en filosofía. Actualmente, es profesor de filosofía práctica en el Departamento de Filosofía de la Universidad de Lund, además, es consultor editorial de *International Journal of Wellbeing* (Nueva Zelanda) y miembro del consejo editorial de *Journal of Applied Ethics and Philosophy* (Japón), también es miembro del Wiley-Blackwell's International. De 2004 a 2010, fue co-editor de la revista *Ethical Theory and Moral Practice* y miembro del Consejo Sueco de Ciencias. En los últimos años, ha publicado diversos artículos sobre la teoría del valor y la meta-ética. Sus publicaciones más importantes son: *Logic, facts and representation: An examination of R. M. Hare's moral philosophy* (1993) y *Personal value* (2011).

PAUL K. MOSER

Doctor en Filosofía. Actualmente, es profesor y director del Departamento de Filosofía de la Universidad de Loyola en Chicago y editor, del *American Philosophical Quarterly*, sus áreas de investigación son: la filosofía analítica, la epistemología y la filosofía de la religión. Es autor de más de 19 obras, sobre epistemología y filosofía de la religión, además de diversos artículos en revistas especializadas. Sus publicaciones más importantes son: *Knowledge and evidence* (1991), *Philosophy after objectivity: Making sense in perspective* (1999), *The Oxford handbook of epistemology* (2002), *Jesus and philosophy: New essays* (2008) y *The evidence for God* (2010).

MICHELE CAPONIGRO

Magíster en Física de la Universidad de Turín. Doctorado en Epistemología de la complejidad, por la Universidad de Bergamo, Doctorado en Física, por la Universidad de Camerino. Actualmente, pertenece al ISHTAR: Indeterministic Sciences and Historico-Philosophical Transdisciplinar Advanced Research CENTER, de la Universidad de Bergamo. Sus áreas de actuación son: interpretación de la mecánica cuántica, filosofía de la mecánica cuántica, cuántica y la naturaleza de la realidad y teoría cuántica. Ha publicado más de una docena de artículos, en revistas especializadas a nivel mundial. Sus publicaciones más importantes son: "Tomograms and the quest for single particle nonlocality", *Journal of Physics* (2007), "Quantum

entanglement and holomovement: An unfragmented epistemology”, *NeuroQuantology Journal* (2009), “Quantum entanglement: can we “see” the Implicate Order? Philosophical speculations”, *NeuroQuantology Journal* (2010), “Quantum interpretation of Vedic theory of mind: An epistemological path and objective reduction of thoughts”, *JCER: Journal of Contemporary European Research Journal* (2010).

ENRICO GIANNETTO

Licenciado en Física Teórica de partículas elementales, por la Universidad de Padua, se especializó, en cosmología de partículas (del universo temprano) en la Scuola Internazionale Superiore di Studi Avanzati: SISSA, de Trieste e historia de la ciencia, en el Domus Galilaeana de Pisa. Doctor en Estructura de la Materia, de la Universidad de Messina. De 1991 a 2001, fue docente e investigador, de la Universidad de Padua, donde enseñó, entre otras cosas: historia de la física y física general. Actualmente, es docente de la Universidad de Bérgamo, donde es Director del Doctorado en Antropología y Epistemología de la Complejidad. También se desempeña, como director del Centro de Investigaciones sobre Complejidad (CERCO) de la Universidad de Bérgamo, además de ser presidente de la Sociedad Italiana de Historiadores de Física y Astronomía: SISFA. Sus áreas de actuación son: historia de la ciencia, la filosofía de la física, la cosmología y la física de partículas. En la actualidad, es editor de la revista *Liberazioni*. Sus publicaciones más importantes son: *Science education and culture conference. The contribution of history and philosophy of science* (2001), *Volta and the history of electricity* (2003), *Un fisico delle origini. Heidegger, la scienza e la natura* (2010), *Note per una metamorfosi* (2011).

THOMAS MORMANN

Doctor en Matemáticas. Lógico y filósofo de la ciencia, ha sido docente en diversas universidades, entre ellas: la Universidad de Viena y la Universidad de Múnich. Actualmente, es profesor del Departamento de Lógica y Filosofía de la Ciencia de la Universidad del País Vasco, sus áreas de investigación son: la filosofía de la ciencia, la ciencia cognitiva, la filosofía de las matemáticas, la lógica, la psicología cognitiva, la topología. Sus publicaciones más importantes son: *Zur Topologie sphärischer Raumformen* (1978), *Argumentieren, begründen, verallgemeinern: zum Beweisen im Mathematikunterricht* (1981), *Struktur und repräsentation, Überlegungen zu einer Konstitutionstheorie der Wirklichkeit* (1995) y *Bertrand Russell, Beck'sche reihe denker* (2007).

DAIAN T. FLÓREZ

Profesora del departamento de filosofía de la Universidad de Caldas. Ha publicado varios artículos sobre filosofía moderna y sobre filosofía de la ciencia y participa activamente de la actividad filosófica nacional. E-mail: daian.florez@ucaldas.edu.co

MACARENA MAREY

Licenciada en Filosofía, por la Universidad de Buenos Aires (2003). Doctora en Filosofía por la Universidad Nacional de La Plata (2010). Actualmente, es docente de las cátedras de Filosofía Política y Estética en la Facultad de Filosofía y Letras, de la UBA. Becaria Postdoctoral del CONICET. Ex-becaria Fundación ANTORCHAS (beca de grado) y CONICET (beca doctoral). Sus áreas de estudio son: la filosofía práctica, la filosofía política y la filosofía kantiana, con especial atención a la *Doctrina del derecho* de la *Metafísica de las costumbres*. Autora de diversos artículos, los cuales han aparecido en reconocidas revistas internacionales, además, de un gran número de capítulos de libros y actas de congresos internacionales, incluyendo los dos últimos congresos kantianos, organizados por la Kant-Gesellschaft.

ÁLVARO REVOLLEDO NOVOA

Magíster en Filosofía y candidato a Doctor en Filosofía. Actualmente, es docente del Departamento de Filosofía de la UNMSM. Sus áreas de estudio son: lógica, teoría del conocimiento, filosofía del lenguaje y filosofía de la psicología. Además, es miembro del Centro de Estudios de Filosofía Analítica (CESFIA), del Perú. Sus publicaciones más importantes son: *En busca de la virtud. La ética del emperador y del esclavo* (2006) y *Menos Platón y más Darwin y otros ensayos* (2009).

MAURICIO VÉLEZ UPEGUI

Licenciado en Español y Literatura, por la Universidad de Medellín, Magíster en Literatura. Actualmente, es docente de la Universidad EAFIT (Colombia). Sus áreas de estudio son: ciencias humanas, lingüística, hermenéutica, filosofía política, historia de la filosofía, filosofía antigua y educación. Sus publicaciones más importantes son: *Los desdoblamientos de la palabra. Variaciones en torno al diálogo* (2005), *El pórtico de Jenófanes* (2006) y *Dar acogida (el motivo de la hospitalidad en la Telemaquía de Homero)* (2011).

CEZAR LUÍS SEIBT

Licenciado en Ciencias de la Religión, por la Pontificia Universidad Católica de Río Grande do Sul (1991), Licenciado en Filosofía, por la Facultad de Filosofía de Nuestra Señora de la Inmaculada Concepción (1996), Licenciado en Psicología, por la Universidad de la Amazonia (2003), Magíster en Filosofía, por la Pontificia Universidad Católica de Río Grande do Sul (2005) y Doctor en Filosofía (conocimiento y lenguaje), por la PUCRS, con una pasantía investigativa en la Universidad Albert-Ludwig de Friburgo, Alemania (2009). Actualmente, es Profesor Adjunto, de la Facultad de Educación de la UFPA y Profesor y Vicepresidente, del Programa de Postgrado en Educación de la UFPA (Belém) Brasil. Sus áreas de estudio son: la filosofía y la educación, con énfasis en antropología, conocimiento, lenguaje y fenomenología hermenéutica. Sus publicaciones más importantes son: "Temporalidade e propriedade em Ser e Tempo de Heidegger", *Revista de Filosofia Aurora* (2010), "Educação, conhecimento e finitude", *Revista Sul-Americana de Filosofia e Educação* (2012), "Heidegger: da fenomenologia reflexiva à fenomenologia Hermenêutica", *Revista Principios* (2012).

PATRICIA MOYA CAÑAS

Licenciada en Filosofía, por la Pontificia Universidad Católica de Chile. Doctora en Filosofía, por la Universidad de Navarra. Actualmente, es docente de teoría del conocimiento y seminario de texto del Instituto de Filosofía de la Universidad de los Andes (Chile), además de ser la Directora del Programa de Fundamentación Filosófica. Sus áreas de estudio son: teorías del concepto, filosofía de la mente, filosofía antigua y medieval, fenomenología, crítica e interpretación. Sus publicaciones más importantes son: *La intencionalidad como elemento explicativo clave de la teoría del conocimiento de Tomás de Aquino* (1999), "El atrevimiento de la razón incitada por la fe en San Anselmo", *Tópicos* (2000), "Fin del mundo, según Tomás de Aquino", *Teología y Vida* (2000), "El internalismo de los estados mentales en John Searle", *Acta Philosophica. Revista internazionale di filosofia* (2003), "Las pasiones en Tomás de Aquino: entre lo natural y lo humano", *Tópicos* (2007), "Filosofía del sentido y crítica del naturalismo en la crisis de las ciencias europeas", *Escritos de Filosofía* (2007).

JAN WOLEŃSKI

Filósofo y Jurista. Doctorado en Jurisprudencia. Actualmente, es docente de University of Management and Information, Rzeszów. Pertenece a la tradición de la filosofía analítica y se considera el heredero de la Lwow-Warsaw School y de la postura lógica de Łukasiewicz. Sus principales

campos de investigación son: la filosofía del lenguaje, la epistemología (principalmente la teoría de la verdad) y la lógica, sin embargo, entre sus numerosos artículos (800) y libros (20), gran parte de ellos, se enfocan en la teoría del derecho y la política. Sus publicaciones más importantes son: *Kotarbinski: Logic, semantics and ontology* (1990), *Alfred Tarski and the Vienna circle: Austro-Polish connections in logical empiricism* (1999), *Essays in the history of logic and logical philosophy* (1999) y *Essays in the philosophy and history of logic and mathematics* (2010).

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 conocidas como MLA (ver ejemplo en la página web de la revista) 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 deben aparecer integradas en el texto, entre paréntesis, así: (apellido del autor número de página). Las referencias completas deben aparecer 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 MLA norms (see an example in the web page of the journal), 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 should appear integrated in the text, in parenthesis, in the following way: (author's last name page number). 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.

Estudios de Filosofía

ISSN 0121-3628

Diciembre de 2011

n^o 44

Contenido

Artículos de investigación

La hermenéutica y los métodos de investigación en ciencias sociales

Darío Alberto Ángel Pérez

La concepción objetiva de la substancia en la Crítica de la razón pura de Kant

Martín Arias-Albisu

Revisiones de la ética de la virtud

Diana Hoyos Valdés

Concepto, palabra y límite: un análisis de las observaciones kantianas referidas al uso e interpretación de términos filosóficos

Ileana P. Beade

Naturaleza del acto liberador del delirio zambrano: acción esencial, evidencia y despertar

José Barrientos Rastrojo

El conjuro de las imágenes: Aby Warburg y la historiografía del alma humana

María del Rosario Acosta López

Modernidad y *parrhesía*. Michel Foucault y la cuestión de la resistencia como *éthos*

Reinaldo Giraldo Díaz

¿Arte o publicidad? Argumentos para defender el carácter artístico de la publicidad

Inmaculada Murcia Serrano

Artefactos técnicos: ¿cuál es el enfoque más adecuado?

Álvaro David Monterroza Ríos

Tiempo cíclico e instante en Cien años de soledad

Carlos Másmela Arroyave

Traducciones

¿Qué es ver?

José Gil/Carlos Vásquez Tamayo

**Instituto
filosofía**

UNIVERSIDAD DE ANTIOQUIA

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



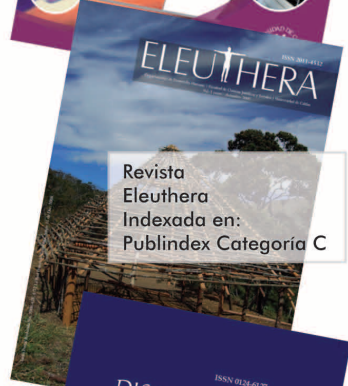
Revista
Agronomía



Revista
Biosalud
Indexada en:
Publindex Categoría B
Lilacs



Revista
Cultura y Droga



Revista
Eleuthera
Indexada en:
Publindex Categoría C



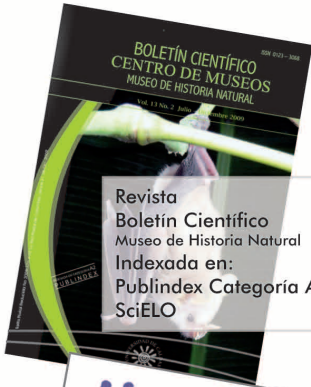
Revista
Luna Azul (On Line)
<http://lunazul.ucaldas.edu.co>
Indexada en:
Publindex Categoría B
Index Copernicus, DOAJ



Revista
Discusiones Filosóficas
Indexada en:
Publindex Categoría A2
Philosopher's Index
SciELO
Ulrich's Periodicals Directory

Revistas





Revista
Boletín Científico
Museo de Historia Natural
Indexada en:
Publindex Categoría A2
SciELO



Revista Colombiana de
las Artes Escénicas



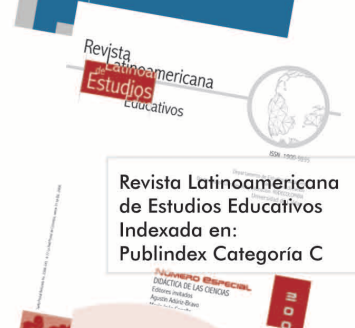
Revista
Veterinaria y Zootecnia
Indexada en:
Publindex Categoría C



Revista
Hacia la Promoción
de la Salud
Indexada en:
Publindex Categoría A2
Lilacs
SciELO



Revista
Jurídicas
Indexada en:
Publindex Categoría C
DialNet



Revista Latinoamericana
de Estudios Educativos
Indexada en:
Publindex Categoría C



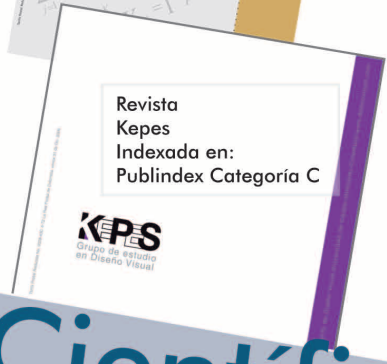
Revista
Vector
Indexada en:
Publindex Categoría C



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



Entregando lo mejor de los **colombianos**



Línea de atención al Cliente Nacional: **01 8000 111 210**

Línea de atención al Cliente Bogotá: **(57-1) 4199299**

► www.4-72.com.co

4°N
72°O

ph

Esta Revista se terminó de imprimir
en julio de 2012 en los talleres de
Capital Graphic
Manizales - Colombia