

## Clasificaciones y géneros naturales: ¿estipulación o descubrimiento?

Alba Amilburu Martínez y Arantza Etxeberria Agiriano  
EHU-UPV, FICE, Departamento de Lógica y Filosofía de la Ciencia, Avenida de  
Tolosa 70, 20080 Donostia -San Sebastián.

La actividad científica establece clasificaciones (explícitas o implícitas) de la realidad al dividirla en géneros o clases (*kinds*) sobre las que interviene experimentalmente o articula teorías. Una discusión frecuente sobre las clasificaciones que se generan al organizar la realidad es si son naturales, o si son, más bien, construcciones que responden a algún tipo de conveniencia. El realismo en las clasificaciones viene a decir que los géneros se descubren, no se estipulan; en ese sentido, representan el orden profundo del mundo.

En este trabajo se examinan algunas cuestiones relacionadas con las clasificaciones del mundo, teniendo en cuenta la tradición de los géneros naturales (GN) examinada por Ian Hacking (1991, 2007), los problemas que aparecen en ella, y la posterior discusión entre la perspectiva esencialista de Kripke-Putnam (1972, 1975) y la de LaPorte (2004). Ambas perspectivas discuten –entre otras cosas– si las clasificaciones responden a un proceso de descubrimiento o son el resultado de alguna estipulación. Aquí este problema se examina atendiendo al caso de la distinción entre lo vivo y lo no vivo (Bedau 2007, Machery forthcoming). Finalmente, se sugiere que muchas clasificaciones, especialmente en la ciencia, requieren algún proceso de iteración entre ambas modalidades.

Cualquier saber requiere de clasificaciones; necesitamos ordenar, organizar y agrupar lo que nos rodea para avanzar en nuestro conocimiento (incluso simplemente para poder hablar de ello). Para Hacking (1991, 2007), los GN son agrupaciones de entidades que responden a un orden natural y se caracterizarían por ser independientes, útiles, únicos y definibles. No obstante, estas características no son suficientes para distinguir entre GN y géneros convencionales o simplemente, géneros no-naturales. Esto le lleva a sugerir que la noción de GN (tal y como se entiende en la amplia literatura que revisa: Mill, Whewell, Venn, Russell, Peirce etc.) es confusa y poco útil, por ello cree que es más conveniente pensar en términos de clases más o menos relevantes. Además, para Hacking los GN no son clasificaciones pensadas para la *contemplación* del mundo sino para *intervenir* en él. En términos de Quine (1969) los GN son categorías adaptativas que han sido y siguen siendo útiles para la supervivencia de la especie aunque estas no siempre coincidan con las unidades o categorías fundamentales que considera la ciencia. Como dice Goodman (1953), los GN se basan en similitudes más o menos intuitivas que no se pueden formalizar.

Filósofos como Kripke o Putnam trataron de solventar esta dificultad en clave esencialista: un término de GN hace referencia directa a una esencia. Esta esencia es la condición de membresía para todo GN. Además, la determinación de esta esencia (condición para determinar la clase) es siempre algo que sólo conocemos a posteriori y por lo tanto es fruto del descubrimiento. Pensemos por ejemplo en el caso del oro. Toda instancia del oro tiene el mismo número atómico, 79. Y la esencia de todo elemento de la tabla periódica es un determinado número atómico. Esta conclusión científica ha sido conocida a posteriori. Sin embargo, la idea de esencia, así como la plantearon Kripke y Putnam, no se ajusta bien al ámbito biológico ya que resulta problemática a la hora de

clasificar la diversidad biológica. No es sencillo hallar o conocer la esencia de una especie determinada, de hecho no contamos con un criterio general (aplicable a todo caso) para clasificar las especies sino que contamos con múltiples criterios distintos. Esto hace que nos cuestionemos la idea de esencia general (condición de membresía para todo GN), al menos en biología (LaPorte 2004).

En este trabajo consideramos que se puede analizar la naturaleza de las clasificaciones relevantes para la ciencia atendiendo a las diferencias entre los criterios estipulados y el descubrimiento de los mismos. Partimos de la idea de que el criterio clasificatorio es siempre el resultado de alguna convención, que puede tanto ser meramente arbitraria (o subjetiva) como estar apoyada y avalada por un estudio empírico. Y al mismo tiempo, por lo menos en ciencia, las categorías ya establecidas siempre dirigen y motivan la investigación hacia el descubrimiento de las propiedades relevantes de las entidades implicadas.

La diferencia entre la estipulación y el descubrimiento puede hacerse patente atendiendo a cómo se piensa sobre la distinción entre lo vivo y lo no-vivo. Mientras que la búsqueda de una *definición de vida* parece tratar de describir teóricamente las condiciones que tiene que cumplir una entidad para que sea considerada viviente, la pregunta sobre si la vida es un GN trata de si la condición de viviente puede descubrirse a posteriori. Por eso decir que la vida no es un GN parece implicar que distinguir entre lo vivo y lo no-vivo depende sólo de la historia o de los intereses humanos (Machery forthcoming), y también que la ciencia no puede ahondar en esa diferencia.

En conclusión, este trabajo trata de examinar algunos de los problemas que se plantean en las clasificaciones, sobre todo en las realizadas por la ciencia. El ejemplo de la distinción entre lo vivo y lo no-vivo podría sugerir que las dos versiones del establecimiento de clasificaciones: la estipulación y la del descubrimiento de propiedades son pasos que la ciencia debe realizar en un proceso iterativo para poder dar cuenta de las categorías utilizadas.

#### Referencias:

- Bedau, M. A. (2007). What is life? In S. Sarkar and A. Plutynski, eds., *A Companion to the Philosophy of Biology*, pp. 455-471. New York: Blackwell.
- Bird, A. (2007). "A Posteriori Knowledge of Natural Kind Essence: A Defence", *Philosophical Topics*, 35.
- Dupré, J., (1993), *The Disorder of Things : Metaphysical Foundations of the Disunity of Science*. Harvard: Harvard University Press.
- Goodman, N., (1978), *Ways of Worldmaking*. Indianapolis: Hackett. -Hacking, I., (1983), *Representing and Intervening*. Cambridge: Cambridge University Press.
- Hacking, I., (1991a), "A Tradition of Natural Kinds", *Philosophical Studies*, 61: 109-126.
- Kripke, S., (1972), "Naming and Necessity". *Semantics of Natural Language*, G. Harman & D. Davidson (eds.), Dordrecht: Reidel, 253-355.
- LaPorte, J., (2004), *Natural Kinds and Conceptual Change*. Cambridge: Cambridge University Press. -Machery, E., (forthcoming) Why I Stopped Worrying about the Definition of Life... And why you Should as Well. *Synthese*.
- Putnam, H., (1975a), "The Meaning of 'Meaning'", *Minnesota Studies in the Philosophy of Science* 7: 215-271. Reprinted in Putnam (1975b).
- Quine, W. V., (1969) *Ontological Relativity and Other Essays*. New York: Columbia University Press.

## **Realismo de perspectivas y causalidad**

Sebastián Álvarez Toledo  
Universidad de Salamanca  
sat@usal.es

(Esta contribución consiste en aplicar algunas de las ideas básicas de R. Giere en *Scientific Perspectivism* a la noción de causalidad)

Una versión de la causalidad ampliamente compartida relaciona la formación de conceptos causales con nuestra condición de agentes intencionales. Hay diferentes propuestas dentro de esta versión, desde aquellas que la creen compatible con una concepción realista de la causalidad (p. e. Woodward) a otras que defienden que la noción de causalidad es meramente subjetiva. Así, por ejemplo, Menzies y Price, quienes identificando la relación entre causa y efecto con la existente entre medios y fines, sostienen que no existen relaciones causales en la naturaleza y que el concepto mismo de causalidad es tan subjetivo y antropocéntrico como puedan serlo las cualidades secundarias. Hay razones para dar por buena esta comparación, aunque no creo que las haya para extraer de ella la consecuencia de que la causalidad es una mera construcción de carácter antropocéntrico. Una cualidad secundaria como el color no es ciertamente sólo una radiación luminosa, no está en las cosas, pero tampoco es un fenómeno subjetivo. El color es el resultado de un proceso neuronal provocado por el reflejo en la retina de una radiación luminosa. Del mismo modo, los ruidos, aunque no son propiamente fenómenos del mundo físico, tampoco son sucesos meramente subjetivos. El árbol que cae en el bosque produce en el aire ondas que se convierten en ruido en el adecuado receptor, y las características de ese ruido dependerán de las capacidades y limitaciones de sistema auditivo de éste. Este carácter dual de las cualidades secundarias es suficiente, según Giere, para afirmar su realidad, aunque se trata de una realidad relativa a una determinada perspectiva. (Este realismo de perspectivas es para Giere la genuina alternativa tanto al realismo objetivista como al constructivismo). Pues bien, si cualidades secundarias como el color y el sonido son el resultado en nosotros de ciertos fenómenos externos como radiaciones electromagnéticas o vibraciones del aire, ¿a qué responde en la realidad nuestro concepto de causalidad? ¿qué relaciones no causales entre sucesos suscitan en nosotros, agentes intencionales, la idea de relación causal? Una buena respuesta a estas preguntas podría consistir en un punto de vista que goza de una larga tradición (p.e, Mach), según la cual, las descripciones causales se refieren en realidad a procesos internos de los sistemas que resultan de formas específicas de correlación entre los valores de sus variables. Desde esta perspectiva, lo importante no es, por ejemplo, que el calentamiento de una masa de gas cause un incremento de su presión, sino el modo en que están coordinados en los gases los valores de temperatura, presión y volumen. Este cambio de perspectiva consiste en situar el elemento activo que desemboca en un fenómeno F (un efecto) de un sistema S, no en los cambios que tienen lugar en el entorno de S (causas), sino en S mismo, en un proceso que tiene lugar en S y al que pertenece el fenómeno F. Podría decirse que, desde este punto de vista, no se atiende al porqué sino al cómo: no se atiende, por ejemplo, a qué causó la muerte por envenenamiento de un animal (la picadura de otro animal, la ingestión de determinado fruto, etc.) sino a cómo ese organismo, ante la presencia de una determinada sustancia en él, reaccionó con un proceso biológico que desembocó en su muerte. Cabría afirmar que la descripción causal de los fenómenos es una descripción de “caja negra”, esto es,

que atiende sólo a correlaciones entre entradas (causas) y salidas (efectos) de un sistema, mientras que en este otro enfoque, centrado en los procesos que tienen lugar en el sistema, las descripciones son de “caja translúcida”, y, por tanto, potencialmente más informativas e incluso más acordes con los objetivos y los métodos de las ciencias. En cualquier caso, se trata de dos perspectivas diferentes, una propiamente causal y otra que se muestra extraña a consideraciones causales.

(Esta diferencia entre una perspectiva causal de los fenómenos y otra no causal puede observarse también en las diferentes interpretaciones existentes del concepto de “capacidad natural”, defendido actualmente por muchos autores como categoría ontológica fundamental. No es lo mismo interpretar tales capacidades como *podere*s *causales*, como cuando decimos que la aspirina tiene la capacidad de aliviar el dolor de cabeza (Cartwright, Bhaskar), que interpretarlas como *disposiciones* de los sistemas para iniciar determinados tipos de procesos internos en circunstancias concretas, como reflejan los conceptos de fragilidad, elasticidad o irascibilidad.) Aceptemos, por tanto, que el concepto de causalidad surge en nosotros debido a nuestra perspectiva de agentes intencionales, pero suscitado por relaciones reales que pueden describirse desde un punto de vista no causal que se muestran en algún sentido más básico y objetivo. De ahí no sigue que no exista la causalidad y que se trate sólo de un concepto puramente antropocéntrico. Continuando con la comparación entre causalidad y cualidades secundarias, podemos concluir que existen realmente relaciones causales en la naturaleza, aunque su realidad es relativa a una determinada perspectiva: nuestra perspectiva de agentes intencionales. Pero, ¿este realismo “perspectivista” de Giere no es un tipo de relativismo? Lo sería si las perspectivas fuesen incompatibles, inconmensurables en sentido de Kuhn, pero no cuando, como es habitual, se trata de puntos de vista parciales y complementarios, entre los que las diferencias son graduales y relativas a la calidad y la cantidad de información que permiten recabar. En nuestra discusión sobre la causalidad los dos puntos de vista comentados son distintos en varios aspectos, pero compatibles y complementarios.

#### Referencias:

R. Giere, *Scientific Perspectivism*. University of Chicago Press, 2006.

Menzies, P. & Price, H. “Causation as a Secondary Quality”. *The British Journal for the Philosophy of Science*, 44, (1993) 187-203.

Woodward, J. F. “Agency and Interventionist Theories”. In H. Beebe, C Hitchcock and P. Menzies (eds.) *The Oxford Handbook of Causation*. OUP, 2009. pp. 234-264.

## El externismo y sus argumentos. Un diálogo con Tim Crane

Silvia Andrés Balsera

[s.andres@filos.ucm.es](mailto:s.andres@filos.ucm.es)

Departamento de Lógica y Filosofía de la Ciencia

Universidad Complutense de Madrid

Nuestro objetivo en esta comunicación es cuestionar la refutación del argumento externista de la Tierra Gemela que Tim Crane ofrece en *Elements of mind*. En concreto, pretendemos mostrar que la exposición de Crane adolece de tres defectos, a saber, (i) este autor exhibe una comprensión poco matizada del externismo, (ii) cabe problematizar su reconstrucción del argumento de la Tierra Gemela en base a la cual propone su réplica internista y, por último, (iii) en su análisis de los pensamientos sin objeto podemos detectar un cierto interés por la semántica de las atribuciones mentales que puede servirnos para diseñar una versión del argumento de la Tierra Gemela que le resulte más difícil de esquivar al propio Crane.

En lo que respecta al primer problema, nuestro autor traza las líneas fundamentales de la polémica internismo vs. externismo ofreciendo una sucinta caracterización de la primera de esas tesis y dejando al lector la tarea de inferir el contenido de la tesis externista en calidad de posición contraria al internismo. Así, según este autor, el internismo mantiene que “un pensamiento acerca de un X podría ser el mismo estado o acto mental incluso si X no existiera”. El externismo, por su parte, defiende una individuación relacional de los contenidos psicológicos al establecer su dependencia respecto de la existencia de su referente externo, lo cual lleva a suponer que los individuos podrían sufrir ilusiones de contenido en el caso de los pensamientos sin objeto. No obstante, creemos que la determinación parcialmente externa de los contenidos psicológicos que defiende el externismo no queda bien caracterizada mediante la etiqueta de “individuación relacional”, puesto que admite varios grados y, además, pensamos que desde esta tesis existen recursos propios para dar cuenta de los pensamientos sin referente (por ejemplo, T. Burge sostiene que estos tendrían contenido proposicional de existir usos lingüísticos asociados al término que expresa el concepto sin correlato externo).

El segundo problema que detectamos en la exposición de Crane tiene que ver con su versión del argumento de la Tierra Gemela. Según este autor, este razonamiento concluye de manera favorable al externismo debido a que en él aparece como premisa el Principio de que el Contenido determina la Referencia (a partir de ahora lo abreviaremos como C-R). Si aceptamos que los contenidos mentales de los gemelos del experimento de Putnam tienen diferentes referencias (algo que Crane admite) entonces C-R le permite concluir al externista que los contenidos mentales de los protagonistas (internamente idénticos) son, sin embargo, diferentes. La estrategia de Crane para escapar a esta conclusión consiste en cuestionar la aplicación generalizada del principio C-R, mostrando que, por ejemplo, no es apropiado para analizar los pensamientos índice. Por nuestra parte no pretendemos defender la legitimidad *in abstracto* de ese principio, sino mostrar que C-R no podría figurar en ninguna reconstrucción razonable del argumento de la Tierra Gemela. En efecto, la utilización externista del escenario gemelo de Putnam ha tenido históricamente un sesgo polémico. Su pretensión es evidenciar que el análisis internista de los contenidos mentales revela insuficiencias y

que, por tanto, existen casos en los que los factores internos a los individuos no bastan por sí solos para fijar sus contenidos psicológicos. Ahora bien, la clave reside en el hecho de que el principio C-R constituye la negación de la tesis internista tal y como la entendía Crane, en cuanto que afirmar (como hace por definición esta tesis) que el pensamiento acerca de X permanece idéntico ante posibles variaciones de X es equivalente a sostener que un cambio en la referencia de los contenidos mentales no implica un cambio en los propios contenidos, con lo que negaríamos C-R. Consecuentemente, si Crane tuviera razón y en un argumento contrario a la individuación internista de los pensamientos apareciera el principio C-R en calidad de premisa (principio que equivale a la negación del internismo) los externistas cometerían petición de principio al emplear frente al internista semejante razonamiento.

El tercer punto del examen de Crane que discutimos parte de su análisis de los pensamientos acerca de objetos no existentes. Según él, la propuesta de B. Russell puede servir de inspiración para el tratamiento internista de este tipo de pensamientos, uno de los cuales podría ser “creo que el rey de Francia es calvo”. El hecho de ver en la forma lógica de ese contenido (cuya expresión es del tipo “el F es G”) una estructura cuantificacional nos capacitaría para dotarle de condiciones de verdad aunque careciera de objeto y, así, establecer su verdad en caso de que existiera un único individuo que satisface la propiedad expresada en la descripción y su falsedad en caso contrario. Mediante este análisis Crane demuestra que los pensamientos que dos sujetos expresan mediante la preferencia “el rey de Francia es calvo” son pensamientos idénticos (aun cuando el valor de verdad efectivo de la expresión fuera diferente en ambos casos) debido al hecho de que las condiciones de verdad de tales pensamientos son idénticas. Ahora bien, esto puede servirnos para reformular de manera más efectiva el argumento de la Tierra Gemela. En efecto, el objetivo de los externistas al utilizar el escenario del experimento mental de Putnam consiste en aportar plausibilidad al hecho de que las condiciones de verdad de los contenidos mentales de los individuos varían al cambiar sus entornos externos. Por ello, al aceptar que en la Tierra Gemela no hay agua (algo que Crane asume al no cuestionar la premisa de que los contenidos de los gemelos tienen distinta referencia), parece difícil admitir que el pensamiento que Óscar gemelo expresa mediante la preferencia “el agua es húmeda” tiene las mismas condiciones de verdad que el pensamiento que Óscar expresa mediante una preferencia homófona. Pero, como hemos visto, el examen de Crane asume que a los contenidos les conciernen esencialmente sus condiciones de verdad, lo cual sugiere que este autor tendría problemas para enfrentarse a un escenario donde las intuiciones hablan a favor de una variación de las condiciones de verdad en función de la variación del entorno.

## Teleosemantics and the Indeterminacy Problem

Marc Artiga Galindo  
Universitat de Girona –Logos  
marc.artiga@udg.edu

The indeterminacy problem is probably the most important objection raised against Teleosemantic accounts. My aim in this paper is to spell out this objection in some detail and consider what teleosemantics can reply to this challenge.

### Teleosemantics

Two important features distinguish teleosemantics from other naturalistic theories of intentionality. The first one is that it analyses content using the notion of *function*. In particular, teleosemantics assumes an etiological definition of function. According to the etiological definition, the function of a trait is the effect that explains why the trait was selected for by natural selection. For instance, the function of hearts is to pump blood (and not to make thump-thump noises) because this is the effect that explains why hearts were selected by natural selection and hence the effect that explains hearts nowadays exist.

The second important feature of teleosemantics is that content is determined by the function of *system* that generates representations. Arguably, the function of representational systems is to produce a mental state when certain external conditions obtain. That is, representational systems were selected by natural selection because they produced mental states M when certain states of affairs S obtained. *Producing M when S obtains* is the effect (a sort of *relational* effect) that explains why representational systems were selected by natural selection and hence it accounts for its existence.

Now, according to teleosemantics, the content of a representation is the state of affairs the mental state is supposed to correlate with (in accordance with the function of the representational system). So in order to find out the content of a representation, we need to look at the state of affairs that has historically correlated with the mental state and whose correlation increased the fitness of the organism.

The classical example used to illustrate this point is the case of leopard frogs (*Rana Pipiens*). Whenever leopard frogs perceive a small black thing moving in a certain direction at a certain velocity, they throw their tongue out and catch whatever they find there. Of course, usually enough these black things are flies, and catching them is what explains that the system that produces mental representations has been fixed in the population. The idea is that the content of the frog's mental state is *there is a fly* because the cases that explain why the representational system was selected by natural selection are cases where this mental state correlated with flies.

### The Indeterminacy problem

We can intuitively formulate the problem as follows: there are many states of affairs that correlated with the mental state M and representing many of them can explain why the representational mechanism was selected for by natural selection. For instance, teleosemantics is compatible with the content of the frog's mental state being *there is a black moving shadow* (instead of *there is a fly*). Since in the environment where frogs evolved, enough black moving shadows were flies, representing both of them can explain why producing such mental states was fitness-enhancing. More generally, the

objection is that teleosemantics warrants multiple content attributions to one mental state in cases where science and common-sense warrant a single content attribution.

The Indeterminacy Problem has been presented under many classifications. I distinguish two versions of the objection: the horizontal and the vertical problem.

### ***Horizontal problem***

The first problem is that there are many different states of affairs in the causal chain leading to the tokening of the mental state that can be said to be the content of the representation. Think again about frogs. Suppose we accept that in the cases that explain why they have this representational mechanism, there was a fly around. The problem is that the mental state is going to correlate with many other states of affairs as well, for instance, a black dot in the frog's retina and a moving black shadow in front of the frog. If representing all of these states of affairs would have increased the fitness of frogs, how can teleosemantics pick up one among all these content attributions?

I think the best strategy for dealing with this version of the problem is Millikan's *consumer-based* teleosemantics (Millikan, 1984, 1993). In short consumer-based teleosemantics claims that what determines content is the function of the *consumer system*, the system that *uses* the representation. In frogs, this system is the tongue snapping mechanism (in contrast to the visual mechanism that *produces* the representation). The idea is that tongue snapping mechanism performs its function successfully only if there are flies (the snapping mechanism does not care as to whether flies are detected as black moving shadows or as red triangles). So this is a way of picking up the right state of affairs that constitutes the content of the representation.

### ***Vertical problem***

However, at this point a new problem arises. Even if we restrict ourselves to the needs of the consumer system, there are many different *descriptions* of the relevant state of affairs that could do the trick. What the consumer system needs are flies, but also food, proteins, a fitness-enhancer, etc... How are we to choose among these contents? No proposal has been able to give a principled way of distinguishing all these contents. Is teleosemantics utterly unable to solve the indeterminacy problem?

My solution comes in two steps. First, instead of trying to rule out some of the descriptions, I claim that frogs represent the conjunction of all of them. That is: the frog's mental state represents *there is a black moving thing, AND there is frog food AND there is a fitness-enhancing thing...* The mental state represents this whole proposition (there need not be parts of the mental state that correspond to parts of the represented state of affairs). In other words, frog's represent the set of possible worlds at which all these descriptions are true. Now, the final twist in my solution consist in pointing out that this content (the conjunction of all descriptions) is roughly coextensional with *there is a fly*. Flies are the kind of thing that is frog food, fitness-enhancing... In other words, the set of possible worlds that make the long conjunction true is approximately the same set of possible worlds in which there are flies. Once we take into account that teleosemantics is a theory about referential content (it seeks to provide the truth-conditions of a given mental state) it follows from this analysis that frogs represent *there is a fly*. I suggest this is how consumer-based teleosemantics can solve the two versions indeterminacy problem.

## **On the plausibility of a similarity-based notion of content**

Stavros Assimakopoulos (stavros@ugr.es)

Departamento de Filosofía I – Universidad de Granada

Traditional semantic theorising has for long been based on the Fregean idea that linguistic communication involves the decoding of the speaker's utterance into an identical proposition in the hearer's mind. The outcome of this process is customarily called the basic proposition expressed by an utterance or 'what is said'. Advances in the field of linguistic pragmatics and, especially the rise of Relevance Theory, which provides the starting point for this talk, have not only pinpointed the fact that "linguistically encoded meaning underdetermines what is said" (Carston 2002:19), but have also provided a detailed account of how non-demonstrative inference manages to enrich the semantic content of an utterance for its basic proposition to be successfully constructed by the hearer.

However, even for relevance theorists, who radically propose that 'what is said' by an utterance cannot be determined solely on the basis of an individual's semantic knowledge, a sentence carries an underspecified semantic representation, which is conceived of as the "common core of meaning shared by every utterance of it" (Sperber & Wilson 1995:9). Lacking more details on what this semantic content materialistically consists in, Relevance Theory's adherence to the Fodorian view of conceptual content seems to bind it with the view that even though basic propositions are only sufficiently similar between interlocutors, semantic contents are identical between them. Even though Fodor acknowledges that "it's old news that the identity conditions for contents, meanings and the like are problematic" (Fodor & Lepore 1999:381), he is prepared to bite the bullet since "the kind of explanations that semantic theories are supposed to give would not survive substituting a similarity-based notion of content for an identity-based notion" (Fodor & Lepore 1999:382). Such a view might be philosophically attractive, but, I contend, goes against the very nature of Relevance Theory, which was originally formulated as a realistic account of cognition and communication aiming at psychological plausibility rather than theoretical convenience.

Against this background, in my talk, I want to focus on the further suggestion that "nobody has the slightest idea how to construct the required notion of content similarity" (ibid), and delineate such an account. One of the most innovative arguments of the relevance-theoretic approach is exactly that we do not need to assume mutual knowledge or any other related substitute which guarantees identity of contexts or propositional contents between interlocutors to explain successful communication. Our cognitive make-up and a powerful inferential mechanism suffice in providing interlocutors with sufficiently similar contexts, basic propositions and implicatures. My proposal is that there is realistically no need for identity at the level of semantic/conceptual content either, since we also have a natural tendency to align our mental contents with the ones of our peers.

Through my talk I want to explore the idea that this co-ordination is the result of various highly specialised cognitive abilities, all hypothesised to be parts of our mindreading capacities, with which we, as humans, are uniquely endowed. In this setting, I will complement the current account of metarepresentation within Relevance Theory with recent research in joint attention, according to which, under normal circumstances, we start attending to stimuli in our environment jointly with a peer, in an attempt to understand his/her intentionality with respect to these stimuli from a very

young age (see, for example, Moore & Dunham 1995). Interestingly enough, our aptitude for joint attention gets all the more fine-tuned while we grow up in corresponding ways with our ability for metarepresentation. Apart from this developmental picture, there is also extensive experimental evidence that an impairment in joint attention, as encountered in certain autistic populations, leads to considerable problems both in the individual's semantic competence and his/her ability for inferential communication.

Apart from the naturalistic account of mental content convergence provided by our mindreading abilities, the present account of semantic content similarity will be supported at a theoretical level through the implementation of Searle's notion of the Background (1983). Given that the Background provides us with the preconditions that enable us to entertain mental representations, it seems reasonable to assume that, if our Backgrounds intersect to some extent, the assumptions that we make and which are affected by our shared Backgrounds, will indeed not be identical to each other, but similar enough for that matter. Since the Background is by definition mental but still linked to our perception of reality in the world (deep Background), like that we walk using our legs, or in a community (local Background), like that we customarily sleep on beds, the dispositions it gives us to entertain certain assumptions about our walking or sleeping in beds will still be the same among all of us in terms of our deep Background or some of us when we share a local Background. Therefore, even if our resulting assumptions about walking or sleeping have different contents due to our different experiences, these contents will still be similar enough to each other in a way that allows for successful communication to take place.

Therefore, since our mental contents have a natural tendency to converge due to mindreading and cross-cutting Backgrounds, this should be in principle enough to suggest that our conceptual contents can correspondingly be not identical but similar enough among individuals. Given, finally, the experimentally supported idea (e.g. Barsalou 1993) that, under normal circumstances, we are able to produce varying representations of the world simulating other people's perspectives in a fast and creative way, the possibility that conceptual contents, much like propositional contents, are customarily inferentially enriched in cases of communication becomes all the more plausible to maintain.

### **References:**

- Barsalou, L. (1993) 'Flexibility, structure, and linguistic vagary in concepts: manifestations of a compositional system of perceptual symbols', in Collins, A. C., Gathercole, S.E. & Conway, M.A. (eds.) *Theories of memory*. London: Lawrence Erlbaum, pp. 29-101.
- Carston, R. (2002) *Thoughts and utterances: the pragmatics of explicit communication*. Oxford: Blackwell.
- Fodor, J.A. & Lepore, E. (1999) 'All at sea in semantic space: Churchland on meaning similarity', in *Journal of Philosophy* 96: 381-403.
- Moore, C. & Dunham, P.H. (eds.) (1995) *Joint Attention: Its Origins and Role in Development*. Hillsdale: Lawrence Erlbaum.
- Searle, J.R. (1983) *Intentionality*. Cambridge: Cambridge University Press.
- Sperber, D. & Wilson, D. (1995) *Relevance: communication and cognition* (2<sup>nd</sup> edition). Oxford: Blackwell.

## Accidents

Sandy Berkovski  
Bilkent University

What kind of events should be considered accidental? This question frequently crops up in the context of history or daily life. Napoleon's flu leads to his defeat at Waterloo. No doubt, we say, it was nothing but an accident. Two strangers find themselves, by a pure accident, on a train, start talking, like each other, and end up getting married within a month.

The view propounded by Poincaré and more recently by Suppes and Ben-Menachem identifies accidentality with instability. A slight change in the initial conditions of the event would lead to a drastically different outcome. When, by contrast, even significant alterations in initial conditions lead to a similar outcome, then we can talk of 'necessities', that is, a class of non-accidental regularities. To put this together:

*Instability theory.* Let  $E$  be an event with the initial conditions  $C_1, \dots, C_n$ . Then  $E$  is accidental if the replacement of  $C_i$  by a qualitatively similar  $C'_i$  leads to  $E'$  which is qualitatively very dissimilar to  $E$ .  $E$  is necessary if the replacement of  $C_i$  by a qualitatively dissimilar  $C'_i$  leads to  $E'$  which is qualitatively similar to  $E$ .

Despite its impressive pedigree, this theory has its weaknesses. I want to mention just one. Suppose that a man stands on the edge of the roof. The wind blows in the fatal direction  $X$ , the man loses his balance and falls to the ground. If the wind did not blow in the direction  $X$ , the man would have lived. His fall seems accidental. But consider: if the man had worn a blue shirt, rather than the red one he actually wore, his fate would have been the same. A small variation in initial conditions does not bring about any drastic changes in the result. Likewise for many other variations. And moreover, some variations plausibly seen as 'significant', would not alter the outcome. Our man could be doing his acrobatics in New York. Would his fate be the same were he to do it in London? We should believe so. But a change in location should be significant. Therefore, for a putatively contingent event  $E$  not every small, or even large, variation would result in a significantly different outcome. Only a small fraction of deviations, presumably the relevant ones, would. But the Instability theory gives us no tools for separating relevant from irrelevant deviations.

Now, the motivation behind the Instability theory has always been in linking accidentality to the impossibility of prediction. When so many minute deviations in the initial setup can effect a major change in the the outcome, the impossibility to know the initial conditions with precision renders virtually any attempt at prediction useless. And I concur: accidents cannot be predicted. But a major point of contention is elsewhere. It is in the possibility of explanation. The defenders of the Instability theory typically argue that accidents can be explained. The failure of prediction, the key characteristic of accidents, does not entail the failure of explanation. While we may be unable to predict whether Federer will beat Nadal, we should still be able to explain Federer's victory after its occurrence.

But I want to argue that accidents cannot be explained. Here it is useful to consider a contrapositive: if  $E$  can be explained, then  $E$  is not an accident. Assuming in the story

above that the meeting of two strangers on the train was an accident, we can still explain why one of them was on that train *X*, and why the other got on *X*. Required details about their itineraries can be supplied. But can we have an explanation why they got on the *same* train? An explanation in terms of *A*'s itinerary would tell why *A* got on the train leaving at 8AM, but it will not tell us why *B* got on the same train as *A*. And similarly, can we explain why people getting on the same train get married? I am not claiming outright that no explanation is available. The idea is rather that, if you begin thinking of an explanation of *E*—even a statistical one—then you thereby begin thinking of *E* as a non-accident.

The suggestion I want, therefore, to make is as follows. Accidents must be identified relative to our background theory *T*. In natural sciences the theory *T* can to a large extent be formalised, whereas in history or daily life it can be equated with a piece of narrative and left at some comfortable level of informality. When an event *E* can be neither predicted, nor explained by *T*, then *E* is accidental relative to *T*.

This proposal has to be strengthened. For sometimes we may have a case when we are unable to explain (or predict) *E*, but nevertheless believe that *E* is not an accident. Perhaps global warming was neither predicted, nor adequately explained. But this is hardly sufficient for bluntly labelling it an accident. I reply that the background assumptions have to be examined more closely. They would probably include a claim that natural phenomena are always governed by laws that would be sufficient for explaining them, even where those laws might forever remain opaque to us. So we have a claim about the existence of explanation, though our ignorance does not permit us to proceed with existential instantiation. As long as we commit ourselves to the claim 'There is an explanation of *E*', *E* will be regarded a non-accident.

## **Autoconocimiento: Autoridad y perspectiva de segunda persona**

Edson Bernd & Manuel de Pinedo (Departamento de Filosofía I, Edificio de Psicología, Campus de Cartuja, 18011 Universidad de Granada)

Aunque varios autores han defendido la necesidad de una perspectiva de segunda persona como constitutiva tanto de la perspectiva de primera como de la de tercera, no conocemos ninguna defensa de la idea de que la autoridad de primera persona presuponga una *autoridad* de segunda persona. En este trabajo queremos argumentar que la capacidad del sujeto de acceder a sus estados mentales no inferencialmente (y de expresarlos por medio de *avowals*) es consecuencia del reconocimiento de una autoridad de segunda persona durante la fase de iniciación en las prácticas lingüísticas. Además, mantendremos que la perspectiva de segunda persona seguirá siendo necesaria a lo largo de la vida del sujeto en la medida en que el autoconocimiento es indisociable de la autocrítica. Que la persistencia de esta perspectiva de segunda persona mantenga el carácter de autoridad de segunda persona es una cuestión que dejaremos abierta.

Con frecuencia el contraste entre lo objetivo y lo subjetivo se establece en términos de una perspectiva de primera persona y una, distanciada, de tercera persona. Ésta distancia, paradigmáticamente, se identifica con la visión proporcionada por la ciencias naturales y, también paradigmáticamente, deja la existencia de perspectivas (de sujetos, de agentes) fuera de su ontología. Esta tensión es un momento fundacional de la filosofía moderna y sigue vertebrando una parte importante de la filosofía contemporánea. Algunos autores (Gomila 2002, Pinedo 2004, Darwall 2006) han sugerido, como vía para superarla, que existe también una perspectiva de segunda persona que es condición de posibilidad de ambas: alguien tiene que enseñarnos, tú a tú, a reconocernos como entidades poseedoras de un punto de vista con respecto a cosas distintas de nosotros mismos, pero también a reconocer que las cosas pueden verse de forma diferente desde otras perspectivas y, así, a ser capaces de modificar la propia. (Autoconocimiento y autocrítica son inseparables, una idea que desarrollaríamos en nuestra ponencia.)

Si puede argumentarse que la posibilidad del conocimiento depende de la capacidad de contrastar puntos de vista, de llegar a algo aceptable desde cualquier perspectiva, el conocimiento de uno mismo no puede ser una excepción. Sin embargo, desde Ryle pocos filósofos han negado la existencia de una asimetría entre el acceso que el sujeto tiene a sus estados mentales y el acceso que tienen los demás (un fenómeno que suele llamarse 'autoridad de primera persona'). ¿Cómo puede el autoconocimiento ser conocimiento si para éste es necesaria la triangulación de perspectivas? ¿Qué perspectivas sobre sí mismo puede triangular el sujeto para tener auto-conocimiento?

La propuesta que tomamos como punto de partida es que la autoridad de primera persona no es una mera consecuencia de la existencia de una perspectiva de primera persona con respecto a uno mismo: de hecho, tal autoridad se nos atribuye por defecto sin preguntar acerca del modo de acceso que hayamos tenido a nuestros estados mentales. Puedo decir tras mirarme en el espejo, a pesar de haber tomado tres cafés y de sentirme energético, que estoy cansado. Quien me escucha ni se plantea preguntarme cómo lo sé, incluso cuando en tal caso cabría dar una respuesta. Según esta propuesta, la autoridad de primera persona es consecuencia de que el sujeto es el único que tiene la capacidad tanto de expresar sus estados no inferencialmente como de atribuírselos observacionalmente, y de contrastar ambas formas de atribución (Borgoni 2009, Bensusan y Pinedo 2007). Pero, además, el acceso observacional a los propios pensamientos, sensaciones, intenciones o emociones pasa a menudo, no por la observación de nuestra apariencia o conducta, sino

por nuestra apertura a tomar en consideración lo que piensan sobre nosotros los que nos rodean.

Si el conocimiento, ya sea del mundo, ya sea de uno mismo, exige el contraste de perspectivas, no puede haber conocimiento antes de distinguir entre uno y el mundo. Como siempre, la fase de transición es especialmente problemática para la filosofía: una criatura que todavía no tiene perspectiva propia que contrastar con ninguna otra ha de ser capaz de resonar, de alguna forma, con la de su enseñante. Nuestra tesis es que, al menos en el desarrollo lingüístico, el aprendiz no tiene más remedio que aceptar la autoridad de segunda persona, de quien le enseña, con respecto a cómo decir qué. La obtención del primer lenguaje es la entrada en el reino de la razón tanto en el sentido de un entrenamiento en los distintos juegos de lenguaje que componen nuestra relación con el mundo, como en los que se expresan estados mentales: el aprendiz no tiene alternativa a la incorporación de las herramientas y procedimientos comunicativos ofrecidos por la enseñante. Sin esa inevitable autoridad de segunda persona uno no podría ser introducido en lenguaje alguno y por tanto no pasaría a formar parte de su comunidad.

Nuestro uso de los pronombres y modos verbales expresa las tres diferentes perspectivas fundamentales para el uso inteligible del idioma. La condición de que esas expresiones en el futuro tengan éxito en el seno de la comunidad lingüística a que se pertenece es que lo exprese dentro de ciertas reglas sintáctico-semántico-pragmáticas definidas por esa misma comunidad. No obstante, dominar el primer lenguaje (convertirse en un usuario competente) incluye la conquista de la autoridad de primera persona. La conquista de dicha autoridad, indicadora de lo que llamamos autoconocimiento, es gradual en un sentido y, en otro, se obtiene de esfera en esfera de conocimiento. Inspirados por la perspectiva ofrecida por Bilgrami (2003) defendemos que existe un marco revelador de esos avances, en grados o esferas, que es la capacidad para la *autocrítica*.

Resumidamente, la autocrítica exige la apertura del sujeto de autoconocimiento a la perspectiva que la segunda persona tiene con respecto al mismo. Bajo ese punto de vista, la autocrítica es la interiorización de la perspectiva de segunda persona para la aplicación de criterios de corrección que van desde la mera auto-atribución de estados intencionales hasta la revisión de creencias y valores. Pero, si la autocrítica es la interiorización de la segunda persona ¿sería ella un mero aparato de persistencia y conservación de la cultura? ¿Cómo se fundamentaría, entonces, la capacidad de responsabilizarse uno mismo con sus deseos, creencias y decisiones?

#### Bibliografía

- Bilgrami, A. 2003, "Edward Said: A personal and intellectual tribute"  
([http://3quarksdaily.blogspot.com/2006/02/akeel\\_bilgrami\\_.html](http://3quarksdaily.blogspot.com/2006/02/akeel_bilgrami_.html))
- Bensusan, H. y M. de Pinedo 2007, "When my own beliefs are not first-personal enough", *Theoria* 58, 35-41.
- Borgoni, C. 2009, *Knowing the nature of one's mind: An externalist basis for self-knowledge*, Tesis doctoral, Universidad de Granada.
- Darwall, S. 2006, *The Second Person Standpoint: Morality, Respect, and Accountability*, Cambridge, Mass., Harvard University Press.
- Gomila, A. 2002, "La perspectiva de segunda persona de la atribución mental", *Azafea. Revista de Filosofía*, 4, pp. 123-38.
- Pinedo, M. de 2004, "De la interpretación radical a la fusión de horizontes: La perspectiva de la segunda persona", en J. J. Acero *et al.*, *El legado de Gadamer*, Granada, Universidad de Granada, pp. 225-235.

## Expresivismo mínimo y autoconocimiento

Cristina Borgoni & Manuel de Pinedo (Departamento de Filosofía I, Edificio de Psicología, Campus de Cartuja, 18005 Universidad de Granada)

El modelo expresivista clásico del autoconocimiento (frecuentemente atribuido a Wittgenstein –IF) sostiene que las autoadscripciones de estados mentales son análogas a expresiones no lingüísticas: “me duele” y el llanto comparten la característica de expresar el estado mental del sujeto que las expresa y, además, ambos carecen de valor de verdad. La crítica inmediata a este modelo es que tanto yo puedo rechazar lo que alguien dice sobre mí utilizando autoadscripciones, como otros pueden usar mis autoadscripciones como elementos para una inferencia. Por esto, los modelos más recientes del expresivismo en autoconocimiento, como los de Bar-On (2004) y de Finkelstein (2008), prescinden de la dicotomía entre expresión y aserción para caracterizar la particularidad de las autoadscripciones. Las autoadscripciones difieren, de este modo, de la atribución de estados mentales a los demás en tanto que estos últimos no son expresiones de estados mentales, sino meros informes sobre ellos.

El objetivo de este trabajo es analizar cómo el expresivismo sobre el autoconocimiento –EA– podría ser fortalecido por una concepción expresivista en semántica –ES–. Para ello, expondremos primero los rasgos característicos de una semántica expresivista, en particular el enfoque desarrollado por Villanueva y Frápolli (2010), llamado “expresivismo mínimo”. A continuación intentaremos aplicar dicho enfoque a dos clases de términos presentes en las autoadscripciones: los términos correspondientes a los estados mentales y el “yo”. Defenderemos que una concepción semántica expresivista de los predicados mentales no sería suficiente para el tratamiento del problema del autoconocimiento porque diluiría la asimetría entre las autoadscripciones y atribuciones de estados mentales a los demás. Sin embargo, con el análisis expresivista del término “yo” sí podríamos dar cuenta de dicha asimetría. Parte de nuestra tarea será pensar sobre qué se expresa con éste término. Nuestro último desafío será explorar una segunda asimetría existente entre la perspectiva de primera y de tercera persona con respecto a uno mismo y nuestra hipótesis es que en este terreno la ES no tendría mucho qué decir.

Villanueva y Frápolli (2010) defienden que un expresivismo mínimo en el campo de la ES tendría que comprometerse con las siguientes tesis:

(1) Ciertos predicados no tienen objetos simples como argumentos, sino complejos de objetos y predicados. Éstos son predicados de ‘segundo orden’. Al menos un elemento de la siguiente lista puede explicarse como predicado de segundo orden: creencia, conocimiento, a priori, pensamiento en primera persona, verdad, bueno, malo.

(2) Los predicados de segundo orden no describen cómo es el mundo.

(3) Los predicados de segundo orden no modifican las condiciones de verdad de las expresiones bajo su alcance.

Si aplicamos este modelo a las autoadscripciones, los predicados mentales presentes en ellas, como por ejemplo “creo que” o “deseo que”, son predicados de segundo orden y por lo tanto: (1) tienen proposiciones como argumentos, (2) no describen el mundo y (3) no modifican las condiciones de verdad de sus argumentos.

La tesis (2) claramente encuentra respaldo en el EA, porque dicho modelo se presenta como rechazo y alternativa al paradigma observacional del autoconocimiento. La tesis (3) dirime el conflicto entre el expresivismo clásico y el actual, en tanto que nos

permite decir que las autoadscripciones no carecen de valor de verdad; los predicados de segundo-orden no modifican las condiciones de verdad de sus argumentos pero sí modifican las circunstancias de evaluación. Sin embargo, los predicados de segundo orden referentes a estados mentales no son de uso exclusivo en las autoadscripciones. Los mismos aparecen en las atribuciones de estados mentales a los demás, por lo que no sería suficiente para entender la especificidad del problema del autoconocimiento.

Una alternativa sería extender el ES hacia el término “yo”. Para ello, tendríamos que modificar la tesis (1). Nuestra propuesta es:

(1') Un predicado puede ser considerado de segundo-orden si funciona como tal: si tiene proposiciones como argumento o si no forma parte de lo representado explícitamente, no siendo parte del significado de la preferencia, pero modificando las circunstancias de evaluación.

Según Anscombe, Lewis y Récanati, el “yo” no forma parte de lo representado explícitamente, pero modifica las circunstancias de evaluación. Si tenemos *p*, yo y yo creo que *p*, tenemos que *p* es lo que determina el estado de cosas que tenemos que contrastar, mientras que yo y yo creo que *p* determinan las circunstancias de evaluación con las que tenemos que contrastar *p*. Otra salida sería entender al “yo” como lo hace Williams (1989), como funcionando como un operador de segundo orden.

Si esta maniobra funciona, tendríamos una explicación conjunta entre el EA y el ES. Sin embargo el elemento crucial en el debate sobre auto-conocimiento que no es todavía explicado es el carácter autoritativo de las autoadscripciones. Nuestra estrategia será hacer uso del fenómeno de inmunidad al error en la identificación del sujeto de la autoadscripción (Shoemaker, Récanati y Lewis) y de la transparencia de lo mental (Evans, Moran). Dicha estrategia da cuerpo además a una cuarta tesis expresivista también prevista por Villanueva y Frápolli: (4) Las expresiones de segundo orden se usan para expresar alguna actitud *A* hacia un contenido específico.

Defenderemos que el doble carácter expresivo de las autoadscripciones expresa la transparencia de lo mental: expresa el compromiso con la verdad sobre el contenido por parte del sujeto, en el caso de la creencia, por ejemplo. Nuestro último desafío será entender cómo este análisis da cuenta de la asimetría entre preferencias sobre uno mismo realizadas desde la perspectiva de primera persona y de tercera persona. Nuestra hipótesis es que el tratamiento de tales casos trascienden el alcance del ES.

### Bibliografía

- Anscombe, E. 1975. “First Person”, en S. Guttenplan (ed.), *Mind and Language*, Oxford, Clarendon, 1975: 45-65.
- Bar-On, D. 2004, *Speaking My Mind*, Oxford, Clarendon Press.
- Evans, G. 1982, *The Varieties of Reference*, New York, Oxford University Press.
- Finkelstein, D.H. 2008, *Expression and the Inner*, Cambridge, Mass., Harvard University Press.
- Villanueva, N. y M. J. Frápolli 2010, “Minimal Expressivism”, manuscrito.
- Lewis, D. 1979, “Attitudes de dicto and de se”, *Philosophical Review* 88, 513-543.
- Moran, R. 2001, *Authority and Estrangement*, Princeton, Princeton University Press.
- Récanati, F. 2007, *Perspectival Thought: a Plea for (Moderate) Relativism*, Oxford, Oxford University Press.
- Shoemaker, S. 1968, “Self-Reference and Self-Awareness”, in Q. Cassam 1994, 80-93.
- Williams, C.J.F. 1989, *What is Identity?* Oxford, Clarendon.
- Wittgenstein, L. 1953, *Philosophical Investigations*, Oxford, Basil Blackwell, 1958.

## The “dual nature” ontology of artifacts” revisited

Fernando Broncano

Departamento de Humanidades: Filosofía, Lenguaje y Literatura.

Universidad Carlos III de Madrid.

[Fernando.Broncano@uc3m.es](mailto:Fernando.Broncano@uc3m.es)

The ontology of artifacts doesn't seem to have been a very much transited trail in contemporary philosophy, but it has recently become the focus of attention: the so-called “Dual Nature Theory of Artifacts” (DNTA) has furnished fresh discussions about the identity and ontological character of artifacts. In the first section of my presentation, I will examine a few questions raised by this approach regarding the ontology of artifacts and then suggest a way to deal with some of them through considering certain properties which concepts and artifacts share. The main idea is that analytical questions about the nature of artifacts cannot be separated from how artifacts help to organize the experience.

The DNTA view defines technical artifacts in this way: “[...] technical artifacts [are] ‘hybrid’ objects that can only be described adequately in a way that somehow combines the physical and intentional conceptualisations of the world” (Kroes, P. A., and Meijers, A., 2006, Introduction: The dual nature of technical artefacts, *Stud. Hist. Phil. Sci.* 37(1), p. 2). The “duality” of DNTA refers mainly to a characterization of artifacts as metaphysically heterogeneous entities: on the one hand, they are material realizations of the designer intentions; and hence they have an intentional character. On the other hand they are physical structures composed of prearranged materials and causal chains.

As far as DNTA borrows certain insights from the tradition of teleological approaches to meaning, it is also supported by the analogy between artifacts and organisms. Such analogy appeals to a shared ontology of an underlying structure, to a teleological character -in particular a functional definition- and finally to an evolutionary history that accounts for both structure and function. However, DNTA has inherited also certain difficulties from the concept of function when it is taken as the basis for developing the ontology of organisms: a) the indetermination between etiological and systemic views of functions, b) the disjunction problem that affects teleological approaches to meaning, that is, the problem of what was the precise referential function a word/artifact was selected for, c) the multiple realization problem that both design and use intentions suffer, d) the current discussions about what a species is, between *phenetist* and *cladistic* approaches, a discussion that concerns how to characterize kinds of artifacts, that is, by a list of shared features or by a common design history.

I will argue hence that the teleological analogy, fruitful as it is, transmits notwithstanding an underlying trouble for DNTA, that is: the assumption that it is possible a functional grasping of necessary and sufficient conditions to determine a particular kind of artifacts. This assumption dismisses the importance of the relational nature of artifacts within the context of a material culture. A deeper reason is that although it is heuristically productive to consider artifacts as analogous to meanings, however, this does not entail that the aspects DNTA focus on are the most relevant ones.

In the second section, I aim to exploit new facets of the hermeneutical analogy between meanings and artifacts as a better way of furnishing a proper ontology for artifacts. Particularly, I will focus on an analogy between artifacts and concepts. However, I will

consider certain contextual and historical properties concepts possess, that is, instead of insisting on the linguistic properties of meanings I will attend to how concepts, as well as artifacts, contribute to organize human experiences.

Firstly, both concepts and artifacts define spaces of possibility. The reason is that they free us from necessity, since artifacts as well as concepts are exercises of spontaneity. In a wide sense, concepts establish what is possible to think. In Sellarsian terms, concepts constitute the “space of reasons”. John McDowell, as tailored to the Kantian suggestion that the spontaneity make us free creatures, explains how it is possible to hold jointly the ideas that the external world rationally constrains our thinking, and that the very act of thinking frees us from the world. External reality, says McDowell, exerts a rational constraint on our thinking because of the receptivity of our experience. We are immersed into reality, but because of the spontaneity our concepts we can move into a new space of possibility.

This capacity to transcend necessity is the most important feature that concepts and artifacts share. An artifact is itself a piece of reality that transcends the facts since it would not exist without the intervention of a creative will that goes beyond the actual world. Moreover, its very existence in a complex network of other artifacts help open new possibilities of reality and new accesses to this reality which would not exist without the presence of such artifact. There is a deep connection between experiences mediated by concepts and actions mediated by artifacts. In both cases the subject is involved into reality by transforming it through her own spontaneity; at the same time, this zone of interaction transforms in turn her subjectivity. That things can be otherwise is at once an exercise of receptivity to external constraints and of spontaneity. It is an effect of being involved into a reality that it is transformed while it transforms us.

Secondly, both concepts and artifacts share a normative character. The normativity of concepts derives from its character of rules. A competent creature that possesses a concept will be able to ask and give reasons concerning the scope of consequences this concept opens. Likewise, an artifact affords a range of actions the agent is able to make and of ends that she could achieve (given the corresponding abilities). These spaces are normatively constituted as far as the uses must be recognized as acceptable by the others participants in the practice, and consequently the agent receives credit by the successful use. Normativity derives in both cases from the character of achievement we attach to uses. Concepts and artifacts share also other features concerning their normative character. For instance, they are discriminative. In the case of concepts, their success conditions allow to determine whether the objects fall or not under the extension of such concepts. This discriminative ability opens the possibility of classifying the world. In a similar sense, an artifact determines a particular opportunity as pertaining to the scope of possibilities the artifact affords. Other shared feature is that both concepts and artifacts must preserve something: content and reference, in the case of concepts, and possibilities in the case of artifacts.

Therefore, the way in which artifacts and concepts allow us to enter into the “Kingdom of the ends” is the same: they situate us so to speak inside and outside of reality. We are inside reality, because concepts and artifacts engage us into a realm of experiences; and we are outside of reality, because they allow us to transcend the experience by dividing the possibilities in those which fall into the range of application and those which do not. In this way, we will hope to explain how to go beyond the biological analogies without jeopardizing the hermeneutical closeness between concepts and artifacts.

## Dependence and the epistemic value of trust

Fernando Broncano.

Departamento de Humanidades: Filosofía, Lenguaje y Literatura.

Universidad Carlos III de Madrid.

[Fernando.Broncano@uc3m.es](mailto:Fernando.Broncano@uc3m.es).

Jesús Vega Encabo

Dept. Lingüística General, Lógica y Filosofía de la Ciencia

Universidad Autónoma de Madrid

Crta. Colmenar, km. 15,4 – 29049 Cantoblanco (Madrid)

[jesus.vega@uam.es](mailto:jesus.vega@uam.es)

Testimony is for us one indispensable source of knowledge. But the fact that in a testimonial exchange there are necessarily involved two intentional agents with their corresponding epistemic states and perspectives motivates skeptical doubts that do not apply to other sources of knowledge. In this paper, we argue that, to the contrary, it is *because* testimony requires the involvement of intentional agents that it acquires its particular epistemological significance.

We propose to view testimony in the light of an interpersonal relation of *dependence*. A relation of epistemic dependence is an asymmetrical relation between the person A and the person B with respect to the knowledge of the proposition p in a certain environment E. In order to define our notion of dependence we introduce first a notion of epistemic perspective that departs in significant ways from current notions in the epistemological literature:

*Epistemic perspective of S*: S has an epistemic perspective on p if and only if S bears a certain relation of cognitive access to p in the environment E, in the circumstances C, and according to the cognitive abilities V such that if S would exercise his/her abilities, he/she would truly believe that p with a certain level of reliability R due to the very exercise of the ability.

Epistemic perspectives can be ordered by their respective degrees of quality according to different epistemological requirements. For instance, one can assess a perspective according to a “safety” condition, that is, according to the possibility of encountering a set of relevant defeaters; or one can assess it in terms of aptness, that is, in terms of the obtaining of truths due to the exercise of those competences that the subjects have, etc. We propose thus an abstract notion of quality of perspective in order to characterize the degree of accessibility the subject possesses in that particular situation where she aims to get knowledge about p.

We can now introduce a notion of *epistemic dependence*:

The *epistemic dependence* of B on A is an asymmetrical relationship established between the perspectives of both agents such that

- a) A has a qualitatively sufficiently good epistemic perspective on p,
- b) B has a poor epistemic perspective on p, and
- c) B can obtain testimonial knowledge on p only by deferring to the epistemic perspective of A.

A situation of epistemic dependence occurs when B is in a position to get knowledge of p conditionally to A’s having knowledge of p and B’s improving her perspective by means of the information provided by A. The resulting concept of dependence allows us

to connect the respective epistemic performances of the agents involved in the testimonial situation by establishing a conditional relation of epistemological success where both A and B play a positive role. The agents' epistemic performance and success depends on the interdependence established between them.

Given these notions, we can characterize a testimonial situation as a situation of epistemic dependence where the link established between both agents is mediated by the mutual recognition they have of being sharing the epistemic quality of their perspectives, and by a positive intentional disposition to link both perspectives such that B is now in a position to access to the truth of p through A's epistemic perspective.

The testimonial situation involves a certain amount of epistemic risk that could lead to an excess of credulity. The particular epistemic risks in testimony derive from the possible lack of sincerity of A when asserting that p or from the difficulties that B can have for assessing the epistemic perspective of A. These very well-known facts have lead many authors either to minimize the epistemic role of the hearer or to require from her a more responsible epistemic behavior when accepting other's words. In our account, we attempt to overcome this alternative. We will argue that, in so far as testimony involves genuine epistemic dependence, both speaker and hearer are required to epistemically contribute to the success of the testimony through a mutual recognition of their respective epistemic perspectives. Part of the epistemological value and significance of testimony lies in the mutual recognition by the epistemic agents of the situation of epistemic dependence they are in.

This view on the nature of testimony gives a certain epistemic value to the attitude of trust exhibited by the receiver of the testimony towards the testifier. So the paper needs to answer some objections that have been raised against the idea that trust attitudes could play any epistemic role in testimonial knowledge. In defense of our view, we take as starting point Keith Lehrer's proposal about the role played by self-trust in the acquisition of knowledge through other sources. We apply this very same model to interpersonal situations. In a recent paper, P. Faulkner (2007) remembers Williams' idea that "trust is a willingness to be dependent". Trust reflects a spontaneous and positive attitude towards a speaker A, by means of which B assesses the situation of epistemic dependence in which both of them are involved. Both these features characteristic of trust –the spontaneity and the recognition of dependence- indicate the exercise of a certain ability to cope with situations of epistemic risk. Trust does not arise in any situation and context: it exhibits a sort of evaluative response to the situation that takes into account a set of parameters implicit in the context, parameters that cannot be defined by the mere presence in the environment of the words by which A states that p. Trust –and this is our thesis- is an attitude through which the epistemic subject establishes a cognitive relation with her environment regarding the possibility of reliably obtaining information from others. In normal conditions, the cognitive agents manifest trust in those contexts where reliable sources of information could be found. So the asymmetrical relations between adults and children or experts and laymen are shaped as epistemically significant in a certain environment prior to the manifestation of trust. A manifestation of trust entails a high-level cognitive integration by an agent that assesses her relation to the environment she is placed in. This is compatible with a minimal degree of learning and conceptual development. Our species is endowed with a sort of calibration system for coping with situations of epistemic dependence.

## Anti-luck Virtue Epistemology: the Whole Story?

Fernando Broncano-Berrocal

Universitat de Girona/LOGOS

ferbroncano@gmail.com

Duncan Pritchard's new theory of knowledge (Pritchard et al., *forthcoming*; Pritchard, *manuscript*) is motivated by two master intuitions which govern our thinking about knowledge. The first one, the *anti-luck intuition*, is the intuition that when one knows one's cognitive success is not a matter of luck. The second one, the *ability intuition*, is the intuition that knowledge requires cognitive ability, in the sense that one's cognitive success is, in a substantial part at least, creditable to one's cognitive abilities. Pritchard's daring point of view is that no theory of knowledge hitherto is sufficient for knowledge. Theories in the market either include (together with the trivial truth and belief conditions) conditions motivated just by the anti-luck intuition or conditions motivated just by the ability intuition (i.e. they include either *anti-luck* or *ability conditions*) and he argues that they are insufficient precisely because no formulation of an anti-luck condition can fully accommodate the ability intuition and no formulation of an ability condition can fully accommodate the anti-luck intuition. According to him, epistemologists have obviated the fact that the two intuitions actually impose *independent* epistemic constraints on our theory of knowledge. For that reason, he proposes a theory which includes conditions of the two types, what grants the independence of constraints:

*Anti-luck Virtue Epistemology (ALVE)*: S knows that p if and only if S's safe true belief that p is the product of her relevant cognitive abilities (such that her safe cognitive success is to a significant degree creditable to her cognitive agency).

First I will remark that in order to achieve a better understanding of the intuitions one should bear in mind that the notion of cognitive success designates two different things:

(T) The *hitting on truth* of the target belief (which is held by the agent).

(B) The *holding* by the agent of the target belief (which is true).

In view of that, I will explain that the anti-luck intuition is concerned only with (T), while the ability intuition both with (T) and (B). Once the intuitions are properly explicated, I will contend that the formulation of ALVE is ambiguous for it gives rise to two possible interpretations:

ALVE-1: S knows that p iff

- i. p is true,
- ii. S believes that p,
- iii. S's reliable cognitive abilities are a relevant significant part of the total set of factors which explain S's *safe* cognitive success.

ALVE-2: S knows that p iff

- i. p is true,
- ii. S believes that p,

- iii. S's belief that p is *safe*,
- iv. S's reliable cognitive abilities are a relevant significant part of the total set of factors which explain S's cognitive success.

ALVE-1 (the one that Pritchard has in mind) differs from ALVE-2 in the following respects: in ALVE-1 the agent's reliable cognitive abilities explain to a significant degree: 1) why the agent *holds* the target belief; 2) why the belief *hits the truth* and 3) why the target belief is *safe*. In ALVE-2 the belief is *circumstantially* safe and the agent's reliable cognitive abilities just explain 1) why the agent *holds* the target belief and 2) why the belief *hits the truth*. My main argument is that ALVE, though better placed than other theories, is also insufficient for knowledge, since it faces the following counterexample:

### *Tempy*

Tempy is an excellent natural temperature detector. She can perceive the temperature of the environment with a margin of error of  $\pm 1^\circ\text{C}$ . As a result, Tempy forms the belief that the temperature in the room is approximately  $n$ , i.e. she believes that  $n$  is within a range of possibly correct values. Tempy also forms the belief that the temperature in the room is  $n$  by consulting a thermometer. Her belief, so formed, could not be easily false (in nearby worlds the temperature is still  $n$  and so indicates the thermometer).

In addition, Tempy has good justification to believe that the thermometer indicates the correct temperature half of the times and exactly  $3^\circ\text{C}$  degrees more the other half. Apart from that, she has no reason for thinking that there is anything amiss with the thermometer. However, the thermometer is in fact broken, and is fluctuating randomly within a given range. Unbeknownst to Tempy, there is an agent hidden in the room who is in control of the thermostat whose job it is to ensure that every time Tempy consults the thermometer the 'reading' on the thermometer corresponds to the temperature in the room.

In the Tempy case (a variation of Pritchard's Temp case) all the conditions of ALVE obtain but Tempy does not know that the temperature is  $n$ : 1) Tempy believes truly that temperature in the room is  $n$ . 2) Tempy *holds* her belief because of her cognitive abilities. 3) The fact that her belief is *safe* and *hits the truth* is explained by Tempy's detecting ability and by the manipulation of the hidden agent.

If Tempy hadn't had the detecting ability, sometimes she would have believed that the temperature indicated by thermometer is correct and sometimes she would have believed that the indicated temperature is  $3^\circ\text{C}$  degrees more than the correct one, so that she would have ended up with false or true unsafe beliefs. But since she displays her detecting ability she is in position to reject the misleading (but grounded) justification, to make her true beliefs safe, to disbelieve false beliefs and to know that the temperature provided by the thermometer is approximately right (i.e. to know that  $n$  is within a range of possibly correct values). Yet she does not know that temperature is  $n$ .

### References

- Pritchard D.H. (*manuscript*) - 'Anti-luck Virtue Epistemology'.  
Url: <[http://www.philosophy.ed.ac.uk/people/full-academic/documents/ALVE\\_000.pdf](http://www.philosophy.ed.ac.uk/people/full-academic/documents/ALVE_000.pdf)>
- Pritchard D.H, Millar, A., Haddock, A. (2010) – Chapter three 'Anti-luck Virtue Epistemology'. *The Nature and Value of Knowledge. Three Investigations*, Oxford University Press: Oxford.

## Endurantism and the temporal grounding problem

Marta Campdelacreu

Universitat de Barcelona & LOGOS

marta\_campdelacreu@ub.edu

The two main theories in the current debate on the nature of the persistence of ordinary objects such as dolphins, trees or rocks are perdurantism and endurantism. Perdurantism claims that objects persist through time by having a temporal part at every time they exist. Endurantism claims that objects are wholly present whenever they exist.

It is often claimed that the temporal grounding problem can be easily dissolved by those versions of perdurantism that are affected by it, but that it is a real difficulty for those versions of endurantism for which it can be posed.

My main objective in this paper is to develop an endurantist answer to the temporal grounding problem.

What is, then, the temporal grounding problem? Let us imagine that an artist creates out of an amorphous amount of clay, at the same time ( $t_1$ ), a piece of clay, Piece, and a statue that we suppose is the statue Goliath. Let us imagine, as well, that later (at  $t_2$ ) the displeased artist remoulds her work into a ball and Goliath ceases to exist. Piece is still there but now it is not coincident with the statue.

The versions of perdurantism and endurantism affected by the temporal grounding problem accept that in cases like these there are two different objects that coincide for a period of time.

Now, the temporal grounding problem is the following. It should be said before progressing that different philosophers offer slightly different lists of *basic/non-basic* properties but that all of the formulations agree with the following statement of the problem.

(1) At the time of coincidence Goliath and Piece share their *basic* properties — examples of shared *basic* properties that are often mentioned are: their material components, their mass, size, shape, at least some of their extrinsic properties, etc.

(2) The *basic* properties of objects determine their *non-basic* properties — examples of *non-basic* properties that are often mentioned are: modal and sortal properties.

(3) Therefore, Goliath and Piece share their *non-basic* properties.

However,

(4) Goliath and Piece do not share their *non-basic* properties. For instance, they do not share their sortal properties (in this paper I focus on these properties).

How can endurantists avoid this paradoxical situation?

The strategy that I will develop on behalf of endurantists begins as the strategy typically used by perdurantists. The perdurantist strategy implies, first, the denial of the intuition behind the temporal grounding problem: at the time of coincidence there are two objects, which share many properties but differ in their sortal properties. After this, they explain the resulting situation by appealing to several features of perdurantist objects that the formulation of the problem would not have taken into account — thus resulting in the alleged paradoxical character of the situation.

Now, let me present the features endurantists claim that objects possess and that explain the situation resulting from the denial of the intuition behind the temporal grounding problem. This will show its non-paradoxical character.

One of the ideas that seem to characterize endurantism is that objects are *wholly present* whenever they exist. I will argue that this means, at least partly, that the sortal property of an object is already determined in the process of its creation, independent of its own futural properties. How this is accomplished depends on the kind of object we are considering and the details will have to be provided by the experts. But, for example, for a statue, it seems plausible to think that the process of its creation—a ‘statue-creation process’, as it can be called—in which its sortal property is determined, involves a certain kind of person with the right type of intentions in the right kind of environment and who makes the right kind of actions in relation to the right kind of material—whatever the case, the details have to be provided by the experts.

Another example: when a piece-of-clay-creation process takes place a new piece of clay, with its sortal property, comes into existence. This process, in my opinion, includes an amount of clay having its parts interconnected in the right kind of way—again however, the details have to be provided by the experts.

Our case is one in which these two different processes of an object’s creation take place at the same time (sharing some of their components even if not all of them) and therefore, from the very beginning, it is determined that there is a statue, Goliath, and a piece of clay, Piece, placed where the processes of their creation determine.

This idea can be combined with the endurantist idea that the persistence conditions of objects are of an *excluding nature*, that is to say, that objects that have persistence conditions which are not compatible are not the same object.

With the two ideas mentioned so far, we can explain how it is that in our situation we have a statue and a different object that is a piece of clay. First, the two different, but overlapping, processes of statue-creation and piece-of-clay-creation constitute the creation, at a given place and time, of a statue, Goliath, and of a piece of clay, Piece. Second, as their persistence conditions are incompatible, Goliath and Piece are two different objects instead of one.

But then, how do we explain that they have so many properties in common? Answer: objects have a material component that can be shared by different objects and that is the source of the shared properties.

In short, an endurantist theory of objects with the following features allows us to give a coherent endurantist account of the situation that results from the denial of the intuition behind the temporal grounding problem: 1) objects have components of a dual nature: a sortal-component and a material component, 2) the sortal component is already determined in the creation process of the object, irrespective of the object’s own futural properties, 3) the persistence conditions of objects are of an excluding nature, 4) the material component can be shared by different objects.

## Warrant and Proper Function

Ariel S. Cecchi

University of Geneva, rue de Candolle 2, 1205 Geneva, Switzerland

This paper deals with the justification of perceptual beliefs. In particular, it concerns external justification of perceptual beliefs as it is proposed by proper function theories. Two different explanations are discussed: Plantinga's proposal of warrant – i.e. which makes the difference between knowledge and mere true belief – in which proper function entails reliability; and Bergmann's proposal of justification in which proper function does not entail reliability. The aim of this paper is to examine the problems of Bergmann's justification in terms of proper function when it is applied to perceptual beliefs. I argue, first, that perceptual beliefs need warrant and it entails reliability. Second, I hold that Bergmann's theory of justification in terms of proper function is insufficient to give an explanation of perceptual justification. I then propose a new account of warrant. Finally, I defend the new proposal and show how it can help to answer some sceptical criticisms addressed to the theory.

The proper function of a system is the function which the system has been designed for. There are two ways in which proper function explanation can be applied: the first is in terms of warrant, as it was proposed by Alvin Plantinga (1993a, 1993b, 1996, 2000); the second is in terms of justification, as it was set by Michael Bergmann.

Plantinga's account of warrant is defined as follows: a subject's (perceptual) belief is warranted if and only if the cognitive faculties producing the belief are (a) functioning properly, (b) true-aimed and (c) such that there is a high objective probability that a belief formed in accordance with those faculties is true. Plantinga also adds another condition, call it the "reliability condition": the cognitive environment in which the belief is produced is *sufficiently* similar to that for which the subject's faculties are designed. According to Plantinga it is not sufficient for warrant that the cognitive faculties function (work) properly, are true-aimed and produce mostly true beliefs in the environments for which they were designed (i.e. that the design plan is a good one). It also requires that the cognitive environment in which our perceptual beliefs is formed is sufficiently similar to that for which the cognitive faculties are designed. This clause is introduced because it can be the case that even if the cognitive faculties are good in the environment for which they were designed, the perceiver could never be in that environment. Other conditions are necessary but we do not need to explain them here.

In a recent book, Michael Bergmann (2006) has proposed a version of justification in terms of proper function that drops the Plantinga's reliability condition. Bergmann's justification is similar to Plantinga's definition of warrant without the clause that the cognitive environment must be sufficiently similar to that for which the cognitive faculties are designed. He claims that the modified account is superior to Plantinga's because it can answer the evil demon objections, the fact that reliability is not necessary for justification. An evil demon's victim could have justified beliefs in spite of the fact that their beliefs are unreliably formed. It is worth to note that Bergmann is concerned with an account of justification rather than warrant. Warrant entails reliability but justification does not. Dropping the reliability clause allows him to speak about justification.

To summarize, Plantinga explains warrant for perceptual beliefs with an account of proper function theory that entails reliability. On the contrary, Bergmann's proposal of perceptual justification involves proper function without reliability.

Having considered these two views, I first of all show that Bergmann's justification in terms of proper function is not an adequate account for the justification of perceptual beliefs. These beliefs need to be warranted and not only justified. I hold that the reliability condition is a necessary clause for warrant in perception. Moreover, Bergmann's account is not grounded on any notion of "function" and that is a problem for his theory.

Second, I outline that perceptual warrant entails two types of proper function processes: perceptual beliefs are warranted, as I distinguish it, by the proper function of a *perceptual* system (or perceptual faculties) and by the proper function of a *psychological* system (or psychological faculties). A true perceptual belief is warranted, on the one hand, if it is the output of the perceptual system's proper function which produces (from the fact of seeing, smelling, hearing, touching or tasting an object) a *correct* perceptual experience. And, on the other hand, if it is the output of the psychological system's proper function that produces a *true* perceptual belief from the (correct) perceptual experience on which this belief is grounded.

I show that Bergmann does not distinguish between these two proper function processes. His account of justification seems to be only given by the proper function of psychological faculties leaving aside the proper function of the perceptual system (the production of correct perceptual experiences). However, it is important to point out that the proper function of psychological faculties alone does not entail the fact that, for instance, our eyes function (work) properly (well) or that the perceptual experience is correctly formed. It rather entails that a true perceptual belief is warranted if it is produced by the psychological faculties' proper function grounded on a properly-formed perceptual experience. Plantinga's account avoids this problem.

Third, I explain the advantages of my account of warrant and its application to perceptual beliefs. I defend the view that warranted perceptual beliefs depend on the proper function of both a perceptual system and cognitive system. What means that a true perceptual belief is warranted if and only if the proper function of a psychological system forms the belief on the basis of a *correct* perceptual experience produced by the proper function of the perceptual system. Finally, I explain how reliabilist accounts of proper function can overtake the problems raised by some sceptical objections like the evil demon scenarios.

## Time Expressions and Reference to Time in Frege

María Cerezo

University of Murcia

Angel d'Ors

University Complutense de Madrid

Gottlob Frege's primary worry in developing his semantic doctrines was to offer an ontologic-semantic justification of his Conceptual Notation, which was expected to be adequate for any science but which he had proposed motivated mainly by his intention to express his theory of numbers. From this perspective, the existence of temporal expressions (conjunctions, adverbs, temporal clauses, verbal tenses and so on) was at the same time natural and alien to Frege's project. It was natural because some scientific claims require to make some kind of reference to time; and it was alien to his project because mathematical truths and, indeed, all truths are according to Frege immutable and atemporal. And it is this latter feature the one that is capital for Frege, and the one that lies behind his ideas. This is the reason why at the last period of the development of his ideas, the one of the *Logical Investigations*, Frege faces a tension between the atemporal nature of thoughts and the placement in time of the speaker who communicates such thoughts.

The debate about temporal expressions in Frege has centred on the account of indexicals that he presents in *The Thought*, focusing mainly on the tension between his theory of sense and the behaviour of indexical expressions. Scholars divide between neo-fregeans who intend to make sense of and develop further Frege's ideas (paradigmatic examples are Evans and McDowell) and direct referentialists who are more worried about indexicality itself than about Frege's motivation. In this paper, we intend to offer an interpretation of Frege's conception of temporal expressions from the perspective of the genesis and evolution of his doctrines from the *Conceptual Notation* up to the *Logical Investigations*.

We start by presenting the fregean doctrines that are relevant to understand his views on temporal expressions, focusing on the relations between sign, sense and reference on the one hand, and the relation of the speaker to each of them on the other hand. In particular, we focus on the following fregean ideas: a) the doctrine of sense as an evolution of the initial notion of mode of determination in the *Conceptual Notation* into the idea of mode of presentation in *On Sense and Reference*; b) the analysis of content in function and argument, and the multiple ways of carrying such an analysis; c) the different relations between a sentence and a sense: even if a sentence can express only its sense, it can also presuppose, indicate or suggest other possible senses.

In the second part of the paper, we pay attention to the explicit occurrences of temporal expressions in Frege's works. In *On Sense and Reference*, Frege deals with adverbial clauses of time that can be used to construct proper names of temporal objects, and with the possibility of quantifying over temporal objects by means of which the sense acquires the generality proper of a law. In *The Thought*, Frege considers the possibility that a sentence indicates something in addition to the sense expressed by means of a temporal adverb, like in "John has not yet arrived". Finally, he addresses the difficulties created by the interaction between the *time of what is said* and the *time at which it is said* (time of utterance), and his recourse to temporal adverbs such as "today", "yesterday", and so on, to account for it. The result is a view in which context plays an

essential role to determine the sense expressed by a sentence, to the point that in cases where a particular temporal determination is relevant and is not linguistically expressed, no thought is expressed without such help.

Some examples of the claims that we make in interpreting Frege's texts are the following. Firstly, Frege's idea that instants and periods of time are objects, and that they can be named and referred to generates an extra multiplicity of ways in which any sentence in which a temporal determination is relevant can be analysed. Secondly, Frege's ideas in *The Thought* resemble his initial idea in the *Conceptual Notation* that a sign can express a mode of determination of the content which is not a mode of presentation. Indeed, we point out that such a possibility is already present in *Considerations on Sense and Reference*, where Frege claims that *the same thought* can be determined as a generalization of an equation between values of functions or as an identity between concepts. In the final part of our paper, we critically review different interpretations offered by other authors. (See the list of References for some examples.)

#### References:

- Burge, T. Sinning against Frege, *Philosophical Review* 88, 1979, 398–432.  
Evans, G. *The Varieties of Reference*, Clarendon Press, 1982. Forbes, G. Indexicals and Intensionality: A Fregean Perspective, *Philosophical Review* 96, 1987, 1-31.  
García V, J.C. Deixis en Frege, *Onomazein* 2, 1997, 305-323.  
Harcourt, E. Frege on “I”, “now”, “today”, and some other linguistic devices, *Synthese* 121, 1999, 329-356. May, R. Frege on Indexicals, *Philosophical Review* 115, 2006, 487-516.  
Perry, J. Frege on Demonstratives, *Philosophical Review* 86, 1977, 474-497.  
Wettstein, H.K. Indexical Reference and Propositional Content, *Philosophical Studies* 36, 1979, 91-100.

## Against Mary with God's Help! – A Novel Approach to Knowledge Arguments

Anna Ciaunica, University of Burgundy - FRANCE - [ciaunica@yahoo.com](mailto:ciaunica@yahoo.com)

Recently, Nagasawa (2008) claimed that both Jackson's Mary and Nagel's bat arguments fail to derive the conclusion that physicalism is false, and drew an interesting parallel between these two knowledge arguments (KA) and two other arguments from philosophy of religion, namely the argument from concept possession and Grim's argument from knowledge *de se*. He consequently argues that the failure of both KA leads us to an interesting metaphysical view, called '*nontheoretical physicalism*', i.e. the position which holds that although the world is entirely physically uniform, it is not necessarily *theoretically* uniform.

In this paper I discuss one particular and unexpected consequence of Nagasawa's '*nontheoretical physicalism*' (NP) for philosophy of mind: more specifically, I argue that its rejection of the thesis that everything in this world is subsumed by *complete* theories of the physical sciences goes much further than Nagasawa's initial intentions, that is to say, NP proves too much and thus undermines the very core of the metaphysical ambitions of today's physicalism itself.

According to Nagasawa, Jackson's formulation of KA can be simplified without changing its soundness as follows: (1) Mary (before her release) is physically omniscient. (2) Mary (before her release) is not omniscient *simpliciter* (because she learns something on her release). Therefore, (3) Physical omniscience is not omniscience *simpliciter*.

Surprisingly enough, Nagasawa says, most philosophers consider as non questionable the idea that:

- (4) In principle, one can be physically omniscient by simply reading black-and-white books and watching black-and-white television.

But in fact, Nagasawa says, what Mary learns by reading text books and watching television is *all theoretically communicable* physical facts. With these background issues settled, Nagasawa formulates his *nontheoretical physicalism* as follows: (i) Physical omniscience is omniscience *simpliciter*. (The physical omniscience thesis is true.) (ii) Theoretically communicable physical omniscience is not physical omniscience (The theoretically communicable physical omniscience thesis is false.) (iii) Omniscience *simpliciter* requires an instantiation of extraordinary epistemic powers to intuit relevant propositions.

The failure of the Mary argument clearly motivates premise (i). But it also motivates premise (ii), namely the idea that we, human beings, including Mary, cannot grasp in an *exhaustive* and *complete theoretical* manner the fundamental features of physical entities and properties. Thus, from statement (ii) one might derive:

- (5) There is *some* physical knowledge which is neither theoretical nor communicable.

My strategy will be to argue that fully developed, NP entails the idea that premise (1) from 'MARY' together with premise (i) from NP are out of reach of *any* human-type knowledge, in contrast to a Godlike-type of knowledge. Hence, the very metaphysical ambitions of physicalism itself are grounded on the *wrong type* of knowledge. The core of my argument might be schematized as follows:

- (6) Physical omniscience *as* omniscience *simpliciter* is an exclusively Godlike-type knowledge.
- (7) Neither Mary (nor Nagasawa) is a Godlike-type creature.

Therefore,

- (8) Neither Mary (nor Nagasawa) can possess physical omniscience *as* omniscience *simpliciter*.

But, as Nagasawa repeatedly urged, the main thesis of physicalism that he wanted to defend, the very one which he explicitly aimed to “save” from Mary’s prison, is premise (i) i.e. that “the physical omniscience thesis *is* true”. Without premise (i), physicalism loses its “metaphysical” aura and becomes a modest empirical theory among many others.

If Nagasawa is right (as I think he is) in claiming that theoretically communicable physical omniscience is *not* physical omniscience *simpliciter*, then we must seriously consider the *other* face of Janus, namely the *theoretically incommunicable* aspect of human-type knowledge. But again, it is very difficult (at least for me) to see how one might develop a *complete* and *exhaustive* explanation of physical reality based on a combination of this communicable / incommunicable “pas de deux” or on a theoretical/nontheoretical one. At this point in the discussion, one might argue that there are as much ‘incommunicabilities’ in the world as there are ‘Marys’.

Who then, except perhaps God, could grasp the entire picture i.e. in both its communicable/incommunicable aspects?

## REFERENCES

- Braddon-Mitchell, D., & R. Nola (eds.) 2009 *Conceptual Analysis and Philosophical Naturalism* (Cambridge Mass.: MIT Press).
- De Caro, M. & Macarthur, D (eds.) 2004: *Naturalism in Question* (Cambridge Mass.: Harvard University Press).
- Nagasawa, Y., 2008: *God and the Phenomenal Consciousness – a Novel Approach to Knowledge Arguments* (Cambridge: Cambridge University Press).
- Smith, Q. and Jolic A. (eds.) 2003: *Consciousness – New Philosophical Perspectives* (Oxford: Clarendon Press).
- Van Gulick, R. 2003a: ‘Maps, Gaps, and Traps’ in Smith, Q. and Jolic A. (eds.) 2003: *Consciousness – New Philosophical Perspectives* (Oxford: Clarendon Press).

## La naturalidad del escepticismo

Jesús Antonio Coll Mármol

IES Manuel Tárrega Escribano

C/ Sancho Panza s/n, San Pedro del Pinatar 30740 (Murcia)

[jesus.coll@um.es](mailto:jesus.coll@um.es)

Thompson Clarke, en su justamente famoso artículo “The Legacy of Scepticism”, hace una pregunta que en la que se muestra lo que se ha convertido en una estrategia de tipo general contra el escepticismo. La idea es: ¿qué está el escéptico examinando, nuestras más profundas convicciones o el producto de un gran fragmento de filosofía sobre el conocimiento empírico realizado con anterioridad a su aparición en escena? Si se trata de lo primero, el escéptico merece una respuesta porque su dictamen implicaría que el conocimiento cotidiano es imposible y ello conllevaría importantes consecuencias prácticas, mientras que, si se trata de lo segundo, nos encontraríamos con que su dictamen haría referencia a una cierta forma de entender filosóficamente el conocimiento que, se sobreentiende, no tendría nada que ver con nuestro modo de entender el conocimiento en la vida cotidiana, dejándolo indemne con sus apreciaciones.

Uno de los campeones de esta estrategia es Michael Williams. En su libro *Unnatural Doubts* desarrolla una estrategia que se acopla perfectamente al esquema de Clarke y declara como artificiosas y filosóficas las dudas que suscita el escéptico contra el conocimiento. Williams distingue en su libro entre dos formas de enfrentarse al escepticismo, a las que denomina posiciones constructivistas y posiciones diagnósticas. El constructivismo intenta responder a los argumentos escépticos en sus propios términos aceptando la legitimidad de las dudas planteadas por los mismos, intentado ofrecer una respuesta satisfactoria a sus retos. Dicha respuesta consiste en ofrecer una teoría positiva del conocimiento que se considera refuta de un modo definitivo a las dudas suscitadas por el escéptico. Williams no cree que el constructivismo sea una respuesta adecuada contra el escepticismo, porque considera que una vez que se acepta la legitimidad de las dudas escépticas el juego tiene un único ganador, y éste no es otro que el escéptico. Sin embargo, esto no representa para Williams una victoria final para el escéptico, sólo el descubrimiento de que una determinada estrategia contra el mismo es infructuosa. Hay en opinión de Williams otra forma de afrontar el escepticismo a la que denomina como diagnóstica. Este tipo de estrategia no acepta la legitimidad de los retos escépticos e intenta mostrar los compromisos lingüísticos o filosóficos asumidos por el escéptico y que convierten en inasumibles sus dudas. Williams distingue entre diagnósticos terapéuticos y teóricos del escepticismo. Los diagnósticos terapéuticos mantienen que los retos escépticos tienen como base importantes errores en el significado que asignan a los términos que utilizan en sus dudas y que los convierten en verdaderos sinsentidos a pesar de su apariencia de ser significativos. Williams no considera que los rompecabezas escépticos carezcan de significado, pues podemos entenderlos sin dificultad –a veces incluso con apariencia de sencillez- y es por eso por lo que prefiere los diagnósticos teóricos del escepticismo. Esta posición no acepta los desafíos escépticos porque están basados en controvertidas ideas filosóficas que están lejos de ser naturales o, en palabras de S. Cavell, *razonables*. Esto significa que el

escéptico necesita para dar inteligibilidad a su desafío de algunas ideas filosóficas que están muy lejos de ser obvias desde el punto de vista de nuestras prácticas epistémicas. Es éste el motivo por el que dichas dudas nos parecen tan poco convincentes a pesar de parecer inevitables cuando se formulan en sus propios términos. Desde esta aproximación los desafíos escépticos no necesitan ser refutados mediante una teoría positiva del conocimiento, sino que podemos oponernos a ellos denunciando las ideas filosóficas que necesitan para cobrar vida. En el caso concreto de Williams éste defiende que lo que denomina como realismo epistemológico y la doctrina de la prioridad de la experiencia son el caldo de cultivo filosófico indispensable del escepticismo. Si tiene razón y esas ideas son como mínimo controvertidas el escéptico no podría siquiera plantear sus retos sin al menos defender previamente dichas ideas, con lo que toda su posición pendería de ellas y se podría mantener una disputa con él más equilibrada que si se considera su posición como arraigada en el más puro sentido común.

La estrategia antiescéptica de Williams, aunque elogiada por autores como Rorty, ha sido respondida asimismo por otros autores como es el caso de Fogelin. Este autor considera que el escepticismo cartesiano en el que se centra Williams ha sido acusado con justicia de ser filosófico y de depender de ideas controvertidas, pero si todo lo que Williams tiene que ofrecer es eso, se trata como mínimo de un camino muy transitado con anterioridad. Fogelin se queja de que Williams sólo considere el caso del escepticismo cartesiano y le acusa de no haber examinado la legitimidad de un enemigo mucho más sutil, interesante y sofisticado, el escepticismo pirrónico. El escepticismo pirrónico, a diferencia del cartesiano, no depende para Fogelin de la doctrina de la prioridad de la experiencia y su modo de enfrentarse al filósofo dogmático está basado en una característica de nuestro concepto cotidiano de conocimiento: la posibilidad de elevar el nivel de escrutinio de cualquier afirmación de conocimiento que hagamos. Este rasgo es absolutamente natural y si el escéptico se basa únicamente en él estaremos condenados a ofrecerle una teoría positiva del conocimiento y entonces –en esto coincide con Williams- la victoria estará del lado pirrónico.

En mi exposición examinaré hasta qué punto el escepticismo pirrónico es realmente un escepticismo libre de elementos filosóficos. Señalaré en primer lugar que el escepticismo pirrónico es un escepticismo acerca de la filosofía, no acerca de nuestras prácticas epistemológicas cotidianas y que, de ese modo, no se trataría de un escepticismo radical. Argumentaré no obstante que este escepticismo puede ser utilizado contra nuestras prácticas epistemológicas cotidianas pero que, en dicho caso este escepticismo volvería de nuevo a ser considerado como antinatural, siendo deudor de la concepción filosófica del conocimiento que se tenga como patrón en dicho juicio. El problema entonces no estará en la falta de naturalidad del escepticismo, sino en la naturalidad o la ausencia de la misma del concepto de conocimiento que manejemos.

## **El vínculo entre Semántica y Pragmática**

Juan José Colomina Almiñana

Universidad de La Laguna / University of Texas at Austin

Durante los últimos treinta años, dentro del ámbito de la Filosofía del Lenguaje, se ha intensificado el debate entre aquellos que pretenden explicar de manera literal el significado del lenguaje (los literalistas o semanticistas) y aquellos que pretenden explicarlo a partir de los elementos contextuales en los que tiene lugar el acto de significación (los contextualistas o pragmatistas). Esta polémica nos permite trazar una diferenciación entre la semántica (el campo de estudio que persigue determinar qué quieren decir las palabras de nuestro lenguaje) y la pragmática (el campo de estudio que busca explicar el uso que los hablantes realizan de las palabras de dichos lenguajes). Por supuesto, esta primera definición establece una dicotomía poco clara, poco útil e imprecisa de estos conceptos.

Para intentar clarificarla, vamos a diferenciar entre tres planteamientos clásicos acerca del vínculo entre semántica y pragmática:

1. La interpretación de una expresión es puramente dependiente de los elementos significativos que la componen sintácticamente. Por lo tanto, las propiedades semánticas tan sólo son propiedades de tipos de expresión, por lo que toda propiedad que una expresión posea de manera relativa a un contexto no será considerada semántica, por lo que su significado variará según el contexto de uso de la expresión. Esto viene a indicar que para poder definir el significado de lo que se dice debemos atender a las propiedades semánticas del lenguaje. (Ejemplos de esta interpretación puramente semántica del vínculo entre semántica y pragmática los encontramos en el planteamiento sintáctico clásico de Montague, pero también en las recientes remodelaciones realizadas por Kaplan, Salmon, Stanley, Szabó, King y otros). [Literalismos o semanticismos]
2. La interpretación de una expresión depende del contexto de su emisión de dos modos: mediante la determinación del contenido referencial de los elementos sensibles al contexto (es decir, a sus condiciones de verdad), conformando el papel semántico del contexto, y mediante la determinación de aquello comunicado a partir del mismo acto de emitir dicho contenido (el papel pragmático). (Ejemplos de esta posición son la teoría de Austin, pero también la de Recanati, Sperber y Wilson, Carston, Bezendehuit y otros). [Contextualismos moderados y radicales]
3. La interpretación de una expresión dependerá de la clase de efectos pragmáticos que afecten al contenido semántico. En este sentido, hay ciertos elementos pragmáticos que tienen efectos sobre dicho contenido pero otros elementos pragmáticos no lo tendrán. (Ejemplos que defienden en cierto modo dicho planteamiento los encontramos en Bach, Perry, Cappelen y Lepore y otros). [Minimalismos y similares]

Como consideramos de obligado recibo para todo aquel que pretenda acercarse críticamente a los estudios sobre el lenguaje ofrecer una explicación de la relación existente entre ambos ámbitos, nuestro objetivo primario es analizar en qué sentido las anteriores posiciones entienden este vínculo y qué implicaciones tienen las mismas dentro de las mencionadas interpretaciones a la hora de dar una explicación de aquello que los hablantes decimos cuando nos expresamos. Para ello debemos tener en cuenta que:

- a) Como indican las críticas a los planteamientos que analizaremos, la semántica por sí sola no parece capaz de determinar el significado de una emisión.
- b) Como indican las críticas a los planteamientos que analizaremos, tampoco parece fácil establecer los criterios bajo los que es posible reconocer qué elementos contextualmente situados son realmente relevantes a la hora de fijar dicho significado.
- c) Pero lo más preocupante de todo, y será nuestro punto de partida aquí, es que, así las cosas, parece que tanto los análisis contextualistas como los semanticistas acaban por defender una imagen irrealista de los intercambios comunicativos al no contemplar al mismo tiempo los elementos contextuales como los semánticos que permiten determinar el significado de nuestras palabras.

Bajo esta mirada crítica, pretendemos aportar una revisión de las dos primeras versiones del vínculo entre semántica y pragmática y abogar por una defensa del tercer posicionamiento, comentando ciertos matices que nos ayudaran a considerar aquellos procesos pragmáticos relevantes a la hora de determinar las condiciones de verdad de nuestras expresiones y que permiten fijar el significado de nuestro lenguaje.

## Sobre la infraestructura procedimental de la interacción comunicativa

Cristina Corredor  
Universidad de Valladolid  
Pza. del Campus s/n, 47011 Valladolid  
[corredor@fyl.uva.es](mailto:corredor@fyl.uva.es)

Esta presentación pretende atender al modo en que la interacción comunicativa puede analizarse como una práctica normativamente regulada, en la que los actos de habla logran un determinado efecto sobre las relaciones interpersonales, dando lugar a nuevos estados en el mundo social y humano de las personas interlocutoras, a través del reconocimiento mutuo de compromisos y responsabilidades, habilitaciones y licencias, expectativas legítimas, etc. Este planteamiento ha encontrado dos desarrollos notables: el análisis del acto de habla de la afirmación en el marco de la pragmática normativa de R. Brandom, y el tratamiento de las fuerzas ilocutivas en la teoría interaccional de M. Sbisà. La teoría de Brandom tiene el enorme interés de ofrecer una explicación del modo en que el contenido semántico se instituye a partir de la práctica discursiva. Pero cabe preguntarse si la importancia fundamental que concede al acto de habla de la afirmación responde a rasgos estructurales de la pragmática normativa que no permitirían extender el mismo análisis a otros actos de habla.

Mi propósito es tratar de poner de manifiesto algunas limitaciones que la pragmática normativa de Brandom puede tener, si se quiere extender a un conjunto más abarcador de actos de habla y sus correspondientes fuerzas. Para ello, tomaré como punto de partida un muy reciente ensayo suyo en el cual ha propuesto algo que es, en mi opinión, del máximo interés para el problema arriba mencionado. Brandom apunta a una dimensión de análisis de la práctica discursiva que podría extenderse desde el acto de habla de la afirmación hasta integrar a otros actos de habla bien tipificados. Esta dimensión es la de las prácticas de reconocimiento recíproco en que se basa la interacción comunicativa. La intuición básica que parece guiar a Brandom es la de que esta interacción puede analizarse en términos de las relaciones normativas (de responsabilidad y autoridad) que las personas interlocutoras se reconocen mutuamente y en las cuales basan su relación.

Mi propuesta es situar el análisis de la interacción comunicativa en el marco de una concepción interaccional de los actos de habla, teniendo en cuenta lo que E.A. Schegloff ha llamado la *infraestructura procedimental de la interacción*. Este enfoque representa una toma de posición crítica ante la corriente quizá dominante hoy día, para la cual los actos de habla tienen el efecto de dar lugar al reconocimiento de las intenciones de quien habla por parte de quien escucha. Se acepta, por tanto, que son las intenciones de la persona hablante las que, preminentemente, determinan el tipo de acto de habla que se lleva a cabo. En contraposición, para la concepción interaccional se hace imprescindible tener en cuenta la participación de la persona interlocutora en la determinación final de la fuerza del acto de habla, así como en el efecto convencional, basado en un reconocimiento mutuo, que ese acto de habla tiene sobre la relación interpersonal y sobre la interacción posterior.

E. Weigand ha contribuido de manera influyente a esta concepción al formular un *principio dialógico* de acuerdo con el cual la unidad mínimamente autónoma de la comunicación no es el acto de habla individual, sino una secuencia que consta, al menos, de dos actos de habla: uno de toma de iniciativa y otro de reacción o respuesta. Con el mismo enfoque, otros análisis sugieren que la secuencia podría necesitar ser mayor para que dé cabida a interacciones en las que se hace preciso reparar o rehabilitar una intersubjetividad que se ha visto interrumpida. En todos los casos, lo que el análisis de casos reales pone de manifiesto es que la fuerza de un acto de habla sólo puede determinarse tras una secuencia de intercambios, y sólo en la medida en que las personas participantes logran llegar a un entendimiento o acuerdo, con frecuencia tácito, respecto a cuál ha sido esa fuerza. Por esto mismo, puede decirse que la intersubjetividad así lograda es procedimental: no se obtiene por aplicación de algún criterio general o sustantivo en relación con el significado de las expresiones empleadas o la eficacia comunicativa pretendida, sino a partir de un proceso de aproximación y negociación entre las personas interlocutoras.

Este planteamiento está estudiando, precisamente, las prácticas de reconocimiento mutuo a que hacía referencia Brandom en su ensayo. Sin embargo, hay un aspecto adicional del análisis de los actos de habla que él parece no haber incorporado, o no de manera explícita y plena, en su propia propuesta. De acuerdo con un análisis fiel a la teoría original de Austin, el reconocimiento de que ha tenido lugar un determinado efecto convencional depende de que se hayan respetado determinadas condiciones para la realización afortunada del correspondiente tipo de acto de habla. Austin defiende la existencia de procedimientos regulados que subyacen al habla y que son los que instituyen determinados efectos típicos o convencionales. Pero éste es el aspecto que no parece haberse integrado en la pragmática normativa de Brandom, donde los compromisos asumidos y las habilitaciones atribuidas no forman parte de las condiciones de realización afortunada de los actos de habla en el sentido de Austin. Lo que se precisa es poder establecer el modo en que alcanzar un estatuto normativo (un compromiso o responsabilidad, una habilitación o posición de autoridad) depende de, o viene conferido o instituido por, el reconocimiento mutuo de las personas interlocutoras. Aunque este enfoque está presente en sus escritos desde el comienzo (cf. su análisis del proceso de *deferencia* de la responsabilidad inferencial), lo que parece tener un peso determinante para su análisis de la práctica discursiva es el perspectivismo de las actitudes normativas. Las actitudes de compromiso o autorización son las que cada participante asume o atribuye, sin que el reconocimiento mutuo parezca llegar a ser constitutivo de los correspondientes estatutos alcanzados en la interacción. Este aspecto representa una dificultad sistemática para la posible extensión de su teoría al conjunto de las fuerzas ilocutivas de los actos de habla.

---

**La información, entre Babel y Pentecostés**  
José María Díaz Nafría y Francisco Salto Alemany  
Universidad de León, Facultad de Filosofía y Letras  
Campus de Vegazana s/n, 24071 León

Las promesas y las pesadillas de la sociedad de la información se encuentran entre los extremos de esas figuras bíblicas. A un lado nos encontramos la tecnofobia y la obscuridad radical del comunicar, al otro la tecnofilia y la transparencia. Y no es que ambas posturas carezcan de buenas razones. Si queremos ver de cerca qué es lo que se entiende por *información* en los muy diversos ámbitos en los que ésta juega un papel relevante (en los medios de comunicación, en las aulas, los talleres, los laboratorios, las bibliotecas o en la calle) y, en particular, en cada una de las disciplinas científicas que la usan de una forma sustancial: nos encontramos la multitud de lenguas del mito babilónico y a pesar de todo una cierta homofonía, un aire de familia que parece traslucir un entendimiento más potencial que fáctico.

Cada una de las comprensiones que se tienen a cerca de los fenómenos que homofónicamente llamamos informacionales atienden a una variedad de realidades y problemas tan dispar que cualquier búsqueda de inter-comprensión no puede hacerse al margen de la problemática a la que dichas comprensiones buscan solución. La propuesta que aquí se presenta consiste en confrontar las diversas visiones de la información dentro del espeso bosque de nociones, metáforas y problemas en los que tales visiones se desenvuelven.

**1. ¿La información, un concepto puente?** En contraste con la incapacidad de aprehensión del concepto de hierro por parte de quienes habitaban en la llamada “edad de hierro”, nosotros en la “edad de la información” sí podemos preguntarnos por la naturaleza de eso que llamamos información y a su vez proponer nuevas formas de comprender su contenido, medida y valor según se estudia y aplica en los más variados contextos científicos, técnicos y prácticos.

Se dice que hay información en células, palabras, antenas, cables, pensamientos, electrones, cerebros, robots, comunidades, bases de datos, instituciones... De modo que hablar de información es algo esencialmente transdisciplinar. Pero es más, por una parte, el concepto ha ganado un papel central en muchas disciplinas dispersando su significado y estableciendo brechas entre ellos (Capurro 2003) y por otra parte, una aproximación teórico-informacional podría establecer nexos de unión entre disciplinas aparentemente irreconciliables, aportando soluciones a algunos enigmas científicos (Lyre 2002). Es decir, a la vez parece potenciar la incomunicación entre disciplinas científicas como acabar con su atávica distancia, y a su vez, el mencionado efecto dispersivo en los significados de la información ha conducido a la creencia de que la información puede ser útil para cualquier cosa, mientras que frecuentemente no es suficiente para afrontar los problemas que pretenden resolverse con su concurso (Brier 2008).

Puesto que estamos apelando a los entresijos de diversas ciencias, una clarificación de los fenómenos informacionales que provea a la vez rendimientos científicos y sociales debe ser, en consecuencia, interdisciplinar (Janich 2008), siendo los principales problemas que habrían de abordarse –si se busca el máximo de comprensión mutua sin reducir la diversidad fenoménica a la que atiende cada punto de vista–: (1) ¿Hay un grupo de conceptos bajo todos estos usos de “información”? (2)

¿Cómo puede garantizarse a la vez la diversidad de los fenómenos informacionales y el rigor de su aprehensión teórica? (3) ¿Puede un concepto refinado de información establecer un nexo de unión entre todos ellos? (4) ¿Cómo pueden preservarse todos los intereses prácticos respecto a la información?

**2. Hacia una aproximación interdisciplinar: BITrum.** El estado del arte en lo que atañe a las teorías de la información incluye muy diversos paradigmas en curso: complejidad (Teoría Algorítmica de la Información, info-computacional); entropía vs. orden (física, química); infónes (teoría de situaciones); sistema (cibernética); evolución (biología, teoría de sistemas), etc. Haciendo que disciplinas diferentes presupongan a menudo paradigmas distintos y de difícil reconciliación (Díaz y Salto 2009).

BITrum constituye una investigación interdisciplinar en el sentido antes propuesto (procurando abarcar el máximo de puntos de vista sin reducción fenoménica) que desde 2008 ha contado con la colaboración de casi una centena de investigadores y alguno de los grupos internacionales constituidos a lo largo de las dos últimas décadas con la finalidad de aproximarse interdisciplinariamente a la información. Para ello se convocan áreas de conocimiento que van desde las telecomunicaciones a la filosofía, desde la biología a la documentación, desde la física a las ciencias sociales; buscando que los diversos puntos de vista respecto a la información puedan aproximarse en torno a problemas comunes y de interés general. Su objetivo es el de lograr una mejor comprensión de la polifacética realidad de la información y un replanteamiento de los problemas aludidos con la intención de encontrar mejores medios para su confrontación.

Con ese propósito general se pretende “elucidar las nociones de información de modo interdisciplinar, buscando definir conceptos máximamente generales sin desatender los intereses y objetivos que alberga cada punto de vista y distinguiendo diferentes ámbitos de análisis: conceptos, metáforas, teorías, consecuencias y aplicaciones” (BITrum 2008). Metodología la investigación se plantea mediante una articulación de grupos de trabajo, en la que la elaboración continua y permanentemente de un “glosario de conceptos, metáforas, teorías y problemas en torno a la información” desempeña un papel central (Díaz, Salto y Pérez-Montoro 2010).

## Referencias

- BITrum (2008). Objetivos (del proyecto BITrum). [En línea]  
<<http://bitrum.unileon.es/proyecto/1-2-objetivos>> [acceso: 10/04/2010]
- BRIER, Søren (2008). *Cybersemiotics: Why information is not enough*. Toronto: University of Toronto Press.
- CAPURRO, Rafael (2003). The concept of Information. *Annual Review of Information Science and Technology*, 37(8), 343-411.
- DIÁZ NAFRÍA, J.M., SALTO ALEMANY, F. y PÉREZ-MONTORO, M. (eds.) (2010). *Glossarium BITri: Glosario de conceptos metáforas, teorías y problemas en torno a la información*. (ed. bilingüe) León: Universidad de León. [En línea] BITrum <<http://glossarium.unileon.es> [acceso: 30/04/2010].
- LYRE, Holger (2002). *Informationstheorie. Eine philosophisch-naturwissenschaftliche Einführung*. Munich: Wilhelm Fink.
- JANICH, P. (2008). *Was ist Information? Kritik einer Legende*. Frankfurt: Suhrkamp.
- DIÁZ NAFRÍA, J.M. y SALTO ALEMANY, F. (2009). Special issue: “What is really information? An interdisciplinary approach.” *TripleC*, 7(2).

## Failures in Translation

Francesca Ervas

Institut Jean Nicod - Paris, Università di Cagliari

[francesca.ervas@gmail.com](mailto:francesca.ervas@gmail.com)

The paper considers three different definitions of translation and the corresponding kinds of “untranslatabilities” arising in the attempt of transposing into another language different levels or aspects of the overall signification of a sentence. According to the first definition of translation, given two sentences  $S$  and  $S^1$ , belonging respectively to two natural languages  $L$  and  $L^1$ , and a certain context  $\{C\}$ :

T-Def1:  $S^1$  is a translation of  $S$ , if  $S^1$  expresses what is linguistically encoded in  $S$ .

It is a quite common experience for translators, facing the problem of transposing what is linguistically encoded in  $S$  into another language, to be forced to choose between two or more sentences which encode in different ways the lexical meaning of the words composing the original sentence. For instance, ambiguity of lexical words is specific of a language and almost never preserved in another language (Kripke 1979). This could be considered a “linguistic accident”, because we are not usually obliged to choose between different lexical items in the target sentence. At any rate, the problem seems much more serious than it could seem. In the process of reference assignment, it is not only a matter of fact that a translator has to choose between different alternatives, losing some aspects of the original sentence, but it is also a “matter of principle”. As it is clearly explained by Tyler Burge in *Self-Reference and Translation* (1978), translation preserves self-reference only if the reference gets lost. Here, translator’s choice becomes necessary, not contingent.

However, translation could have a much weaker sense, as if it were the practice of matching what is communicated through two sentences belonging to different languages. Given a certain context  $\{C\}$ ,

Translation (T-Def2):  $S^1$  is a translation of  $S$ , if  $S^1$  expresses what is communicated through  $S$ .

According to this weaker sense of translation, it is always possible to paraphrase or reformulate the meaning of a sentence through a sentence of another language, in order to communicate a certain thought. However, if the first definition of translation was too strong, the second one seems too weak, ending up by including translations which depart too much from the original sentence and that we would not recognize as appropriate translation (Ervas 2009). In fact, what is communicated by an utterance in a given context almost never coincides with what is linguistically encoded in it and sometimes includes what is not directly part of what the speaker has actually said (or written). But how much could a translator depart from what the speaker has said? There seems to be a gap between the translation of what is linguistically encoded and the translation of what is communicated that is worth to be explored, in order to better understand what exactly a translation is.

Jerrold Katz, in *Effability and Translation* (1978), argued that the essential translatability between languages is based on the “Effability Principle” (EP), according to which:

(EP) Each proposition can be expressed by some sentence in any natural language.

It could seem quite trivial if we interpret it in line with the weaker sense of translation proposed in Def2 and we mean “communicated” by the word “expressed”. However, in his *Some Logical Problems of Translation* (1978), Edward Keenan reformulated it in a much stronger sense in his “Effability Hypothesis” (EH):

(EH) Anything that can be thought can be said [in one natural language].

However, as it has been shown (Carston 2002), “what is said” could diverge not only from “what is linguistically encoded” (Def1), but also from “what is communicated” (Def2). Hence, we should assess a third notion of translation in line with (EH): given a certain context {C},

Translation (T-Def3):  $S^1$  is a translation of S, if  $S^1$  expresses what is said through S.

At any rate, if translation process requires a set of communicative assumptions and intentions of the speaker in order to translate a minimal notion of “what is said” by S, what will be required to translate “what is communicated” by S? The paper discusses the role of context {C} in each definition of translation (Perry 1997) and try to answer this question, showing that T-Def2 best fits translators’ real practice, making T-Def3 useless and superfluous. On one hand, analysing different cases of “local untranslatabilities”, it seems clear that T-Def3 could be so widened that it finishes to coincide with T-Def2, the translation of “what is communicated” through S. On the other hand, an explanation of local untranslatabilities involving T-Def2 could comprise the data explored through T-Def3 and would surely be more economical, giving a unitary account of the data.

## References

- Burge T., *Self-Reference and Translation*, in F. Guenther - M. Guenther-Reutter (eds.), *Meaning and Translation. Philosophical and Linguistic Approaches*, Duckworth, London, 1978, pp. 137-153.
- Carston R., *Thoughts and Utterances. The Pragmatics of Explicit Communication*, Blackwell, Oxford, 2002.
- Ervas F., *Defining Translation: Semantic Equivalence versus Functional Equivalence*, paper presented at the Congress *Translation and Philosophy*, Carleton University, Ottawa, May 23-25, 2009.
- Katz J.J., *Effability and Translation*, in F. Guenther - M. Guenther-Reutter (eds.), *Meaning and Translation. Philosophical and Linguistic Approaches*, Duckworth, London, 1978, pp. 191-234.
- Keenan E., *Some Logical Problems of Translation*, in F. Guenther - M. Guenther-Reutter (eds.), *Meaning and Translation. Philosophical and Linguistic Approaches*, Duckworth, London, 1978, pp. 157-189.
- Kripke S., *Speaker's Reference and Semantic Reference*, in P.A. French, T.E. Uehling - H.K. Wettstein (eds.), *Contemporary Perspectives in the Philosophy of Language*, University of Minneapolis Press, Minneapolis, 1979, pp. 6-27.
- Perry J., *Indexicals and Demonstratives*, in R. Hale - C. Wright (eds.), *Companion to the Philosophy of Language*, Blackwell, Oxford, 1997, pp. 586-612.

## **Gravity Probe-B Data Analysis: A case for severity to replace use-novelty criteria and the dangers of relying on scientists' current epistemological intuitions**

Christopher Evans

Depto. de Filosofía, Universidad Autónoma de Barcelona, 08193 Bellaterra, Barcelona.

[ToffaEvans@GMail.com](mailto:ToffaEvans@GMail.com)

This paper builds on previous work which has considered the inappropriacy of much of standard analytical epistemology (SAE) for both describing justification and warrant, and prescribing practice in experimental natural science. In it, I endorse a severity criterion that should override use-novelty criteria for justifying tests of claims that are the target of experimentation, in relation to the counterfactual relationships that hold between data and the phenomena they are seen as evidence for. I depict the case of disagreements concerning the Gravity Probe-B (GPB) data analysis as one that shows that the prevailing epistemological attitudes among scientists are insufficient as a basis for decision making and may lead to over-hasty conclusions that lack epistemological rigour. The paper thus uses the GPB case to support calls for promoting a revised approach to epistemology in the philosophy of science.

SAE typically reflects on the relationship of support offered to knowledge claims through the status of the evidence in favour of such claims or the reliability of the processes involved in generating that support. However, SAE accounts are concerned with the different logics of the evidential relationship (and principally with evidence arising through perception). This approach does not capture how experimental scientists typically evaluate evidence; not through logical relationships but through empirical discovery of actual difference making causes together with data analysis and interpretation that often involves complex statistics. I favour the scheme put forward by Bogan and Woodward (1988) to address scientific evaluation of evidence, in which data is regarded as evidence for phenomena. The evidential support relationship between data and phenomena seems to be captured by the counterfactual relationships in which data stand to the propositions that are the target of investigation. Experiments are designed, and datasets are manipulated and interpreted in order to arrive at the necessary counterfactual sensitivity for target claims to be supported or refuted.

Two related worries concerning the modelling and interpretation of data are: non-novel use of data, or double-counting; and the lack of severity of tests of models. Many authors have defended some form of a “novel use” criterion for data in an attempt to disallow the re-use of data that were used once in the construction of specific hypotheses, as evidence in support of those same hypotheses. Such criteria capture the intuition that prediction (of novel results) provides a more stringent test of a hypothesis than accommodation (of pre-existing data)—which logical analyses of the evidential relationship between hypothesis and evidence fail to do—which seems to be common among the scientific community. Mayo (1991; 2008) has defended the idea that this intuition is only a partial expression of a broader epistemological requirement for severity in the tests that we use to test hypotheses experimentally. Although on occasion use-novelty may coincide with a severity requirement, there are important ways in which non-novel uses of data can provide severe tests of hypotheses and also novel uses may fail to meet standards of severity. Thus Mayo explicitly defends double-counting in certain scenarios, in favour of ensuring the severity of the tests faced by hypotheses constructed and tested using the same dataset.

I see the analysis of the counterfactual relationships between experimental evidence and target hypotheses combined with a probabilistic severity criterion of the type advocated by Mayo as the best way to both describe the reality of scientific experimentation in a useful way that is meaningful to scientists and reflects their concerns, and to set out clear aims and goals for test passing that reflect epistemological requirements but avoid over-simplistic or rigid criteria that scientists often seem to adopt. The dispute that arose in 2008 between NASA and scientists working at Stanford University over continued funding for GPB is an ideal case to illustrate how such an approach can help to clarify issues of epistemic justification and empirical warrant, and thereby to resolve such arguments. As it was, the NASA Astrophysics Senior Review Committee decided to discontinue funding of the GPB data analysis before final results were published and cited in its report what it considered to be well-justified scepticism in the scientific community as one of the reasons.

GPB was designed and executed to produce datasets,  $D_x$ , from measurement procedures that could reliably be used to support the claim that, over a given period of time, the spin axis of a gyroscope in free fall around the Earth had changed by amount  $x$  in a given direction: had  $x$  drift not occurred,  $D_x$  would not have been produced. Thus the experimental data could be offered as evidence in favour of the phenomenon of interest. (I will not here consider the relationship between this phenomenon and the related theory claims involving a given  $x$  constraining certain parameters that are used within a particular theoretical framework to characterise spacetime structure and thereby tightening existing constraints on theory choice which could be seen as evidence supporting the predictions of General Relativity.) The actual datasets produced in this one-off experiment bore little relationship to those expected. Since altering data production was not an option, the GPB team embarked on a long and complicated process of interpretation to identify and isolate unwanted components in the data, and separate them to leave the target signal appreciable.

Clearly this is a case in which supporting evidence fails the criterion of use novelty: the same dataset was used to hypothesise the causes of disruption, to model the unwanted effects, and then as evidence for a certain drift occurring in the gyroscopes once the unwanted effects had been removed from the data. For many scientists this double-counting is clear justification for scepticism regarding the validity of results and therefore supports NASA's claim of scepticism "in the scientific community". However, this does not mean that the propositions defended at the different stages in the data analysis would not have passed a severity criterion, despite failing the use-novelty criterion. By identifying the counterfactual relationships between data and target propositions at the key stages in the data analysis, I aim to show that at each stage the sensitivity of these relationships is sufficient in terms of meeting a probabilistic severity criterion, to support the target claims. Thus the NASA decision to discontinue funding may have been hasty based on prevailing attitudes in the scientific community—which may be over-simplistic and benefit from renewed vigour in pursuing reformed epistemological criteria—rather than epistemological rigour.

**References:** Bogan J. and Woodward J. (1988): "Saving the Phenomena", *The Philosophical Review*, **97** (3), pp. 303-352.

Mayo D. G. (1991): "Novel Evidence and Severe Tests", *Philosophy of Science*, **58** (4), pp. 523-552.

Mayo D. G. (2008): "How to Discount Double-Counting When It Counts: Some Clarifications", *British Journal for the Philosophy of Science*, **59** (4), pp. 857-879.

## Self-Deception and Self-Knowledge

Jordi Fernández  
Philosophy Department  
University of Adelaide

The purpose of this talk is to explain a type of self-deception based on a certain model of self-knowledge. To illustrate the relevant type of self-deception, consider the following vignette:

Lately Jack has been avoiding reading any magazine or newspaper article on medical issues. If they come up in a conversation to which he is a party, he changes the topic. He has been scheduled to have a regular check-up with his doctor several times, and he forgets about it each time the appointment is scheduled. Eventually, Jack's relatives have asked him whether he believes that he is sick. But Jack sincerely denies believing that.

There are two remarkable facts about this kind of case:

- (a) There is a tension between the subject's speech and the subject's actions. Jack behaves as if he believed that he is sick. Except for the fact that he claims not to believe that he is sick. Call this tension the 'conflict of' self-deception.
- (b) We also seem to think that Jack is, in some sense, at fault while being self-deceived. Judging a subject to be self-deceived amounts to a *criticism* of that subject. Call this feature of self-deception the 'normativity of' self-deception.

The object of our investigation will be self-deception as characterized by its conflict and normativity.

A natural hypothesis about self-deception is that, in self-deception, you intentionally get yourself to believe something that you yourself believe to be false. Let us call this approach 'intentionalism.' Intentionalism runs into the so-called 'static paradox' and 'dynamic paradox' of self-deception. The static paradox arises from the fact that, according to intentionalism, the self-deceived subject must hold two contradictory beliefs at the same time, which seems *prima facie* impossible. The dynamic paradox arises from the fact that, according to intentionalism, the self-deceived subject must intend to get herself to believe something that she thinks to be false. But it is hard to see how the intention in question could be carried out successfully.

There is an alternative approach to self-deception according to which the self-deceived subject forms a false belief due to the causal influence of a motivational state, typically a desire. Let us call it 'motivationalism.' Motivationalism avoids the two paradoxes of self-deception, but the motivationalist has trouble accounting for the conflict of self-deception. Suppose that Jack wants not to be sick and, as a result, he disregards or misinterprets the evidence that suggests that he is sick. What explains the fact that he avoids being exposed to information about medical issues and he avoids seeing his doctor? Certainly not his desire; you would expect someone who wants not to be sick to be very interested in finding out whether he is sick or not. Thus, it seems that if we endorse motivationalism and attribute to Jack a desire not to be sick, then we will have trouble fully explaining the conflict of his self-deception.

The guiding thought behind the account of self-deception to be proposed is that self-deception is a failure of self-knowledge that takes place when we misuse our capacity for it. Thus, this account of self-deception will rest on a model of how we acquire knowledge of our own mental states, specifically our beliefs. The account is essentially the following. I suggest that, in order to self-ascribe beliefs, we usually look, as it were, past them. Normally, we form beliefs about the beliefs that we have based on grounds that we have for those beliefs themselves. Thus, if I believe that I have the belief that I am sick, then I have formed that higher-order belief on the basis of (what I take to be) evidence for believing that I am sick. I will use 'bypass' to refer to this procedure. We can summarize the proposal thus:

The Bypass model

For any proposition P and any subject S:

Normally, if S believes that she believes that P, then there is a state E such that

- (i) S's meta-belief has been formed on the basis of her being in E.
- (ii) E constitutes grounds for the belief that P in S.

Self-deception should be construed as a failure of self-knowledge of a particular kind. It is the kind of failure that occurs when one not only commits a mistake, but the mistake is also the result of one's negligence.

Consider Jack. On the one hand, he behaves as if he believed that he is sick. On the other hand, he claims not to believe that. This conflict can be accounted for if we take Jack to be making a mistake about his own beliefs. His actions can be explained as expressions of a first-order belief, namely, the belief that he is sick. And his claim that he does not have that belief can be explained as an expression of a higher-order belief, that is, the belief that he does not believe that he is sick.

What about the normativity of self-deception? The reason why we have the intuition that self-deception is objectionable is that we sense that the subject is violating a certain epistemic norm. One should not form a belief in the face of evidence against it unless there are overriding considerations in support of that belief. There is an interesting instance of this general norm: For any proposition Q, one should not believe that one does not believe that Q if one has adequate support for believing that one believes that Q unless there are overriding considerations in support of believing that one lacks the belief that Q. If the explanation of the conflict of Jack's self-deception is correct, Jack believes that he is sick. So Jack must have grounds for believing that he is sick. Now, if the bypass model is correct, whatever grounds Jack has for his belief that he is sick constitute, in Jack, adequate support for the belief that he believes that he is sick. So he has adequate support for believing that he believes that he is sick. But Jack actually believes that he lacks the belief that he is sick. It is no wonder, then, that we have the intuition that his condition is objectionable. After all, Jack is being blind to the fact that he has adequate support for a meta-belief that is contrary to the meta-belief which he has actually formed. Clearly, for Jack to hold on to his meta-belief in such a situation is for him to behave in an epistemically negligent way.

## Autoridad de la primera persona y conocimiento de las propias acciones

Martin Fricke

CEPHCIS, UNAM, Mérida, Yucatán, México

En su libro *Authority and Estrangement* (Moran 2001) Richard Moran afirma una analogía entre nuestro conocimiento de las propias creencias y el de las propias acciones. En esta ponencia exploraré la función y eficacia de la analogía.

Moran menciona esta analogía cuando trata de explicar la “transparencia” entre las preguntas “¿Creo que  $p$ ?” y “¿Es verdad que  $p$ ?”. A menudo, cuando nos preguntan si creemos que  $p$  no examinamos nuestras creencias sino nos preguntamos si es verdad que  $p$  o no. La “pregunta de la primera persona en el tiempo presente sobre la propia creencia se contesta con referencia a (o consideración de) las mismas razones que justificarían una respuesta a la pregunta correspondiente sobre el mundo.” (Moran 2001, 62, cf. también Evans 1982, 225.) Moran se pregunta cómo tal transparencia es posible y en sus intentos de dar una respuesta suele mencionar el conocimiento que tenemos de las propias acciones. Sugiere que aquí también es posible llegar a un conocimiento sobre mí a partir de un razonamiento sobre un asunto diferente. La pregunta “¿Qué es lo que haré?” se contesta a partir de una reflexión de la pregunta “¿Cuál es la acción más atractiva/ razonable/ o mejor (en cualquier otro sentido)?”.

¿Por qué nos refiere Moran al caso del conocimiento de las propias acciones? Una posibilidad es que saber de las propias acciones nos permite saber de las propias creencias. Muy simplificada: Sé qué es lo que hago. Ya que yo soy el que *hace* mis creencias, también sé cuáles creencias tengo. En una teoría de esta forma, el conocimiento de las propias acciones es una precondition para conocer las propias creencias. Carlos Moya parece interpretar Moran según este esquema (cf. Moya 2007). Basándose en algunas observaciones que Moran hace acerca de la teoría de la acción de Elizabeth Anscombe, Moya sugiere lo siguiente: Si una acción intencional es el resultado de una deliberación entonces el agente debe conocer su acción porque en la deliberación se considera la “razón primaria” de la acción. Esta “razón primaria” (un concepto de Davidson) incluye una descripción de la acción. Según Moya se sigue que si el agente conoce la razón primaria de su acción (y debe conocerla si la acción es el resultado de una deliberación) entonces sabe qué acción realiza o realizará. Ahora bien, si la formación de una creencia también es una acción intencional, entonces el agente también debería saber, por las mismas razones, cuál es la creencia que está formando. Moya argumenta convincentemente que tal explicación de nuestro conocimiento de las propias creencias es circular.

Pero quizá la interpretación de Moya no sea correcta. Tal vez cuando Moran alude al conocimiento de las propias acciones no quiera dar el elemento faltante en su explicación del conocimiento de las propias creencias. Más bien, sólo da una analogía. En el caso del conocimiento de las propias acciones simplemente es más fácil ver por qué las dos preguntas anteriormente mencionadas son “transparentes” en el sentido de Moran. Éste entonces trata de “transferir” la idea de transparencia al conocimiento de las propias creencias.

Me parece que esta segunda interpretación es correcta. El ejemplo del conocimiento de las propias acciones sólo tiene fines de ilustración, no provee un elemento adicional o un caso más fundamental en la explicación de la autoatribución de creencias. Sin embargo, quizá Moya tenga razón cuando supone que algo falta en la teoría de Moran. Este mismo dice: “Supongamos que estás tratando de contestar la

pregunta de si P, y entonces deliberas hasta la respuesta de que sí, de hecho, P; nada en mi teoría explica cómo llegamos de este conocimiento de que P al autoconocimiento de que *creo* que P.” (Moran, 2003, 409.) Pero quizá, para explicar la autoridad epistémica que tenemos en el conocimiento de las propias creencias, debemos decir algo sobre cómo llegar de la creencia de que P al autoconocimiento de que *creo* que P. Así quizá también, para explicar la autoridad que tenemos en el conocimiento de las propias acciones, debemos decir algo sobre cómo llegar de mi decisión de que  $\phi$  es la acción que debería hacer al conocimiento de que  $\phi$  es la acción que realizaré. En las palabras de Shoemaker [véase Shoemaker 2003, 400], Moran no investiga la *metafísica* del autoconocimiento. No le interesa la pregunta de cuáles son los mecanismos que permiten que una consideración de cuestiones sobre el mundo produzca un conocimiento de la propia mente. Más bien, Moran sólo se ocupa de la pregunta epistemológica: ¿Qué legitima la transparencia? ¿Pero es posible tratar la cuestión epistemológica sin considerar también la pregunta metafísica?

### **Bibliografía**

Evans, G. (1982), *The Varieties of Reference*, Oxford, Clarendon.

Moran, R. (2001), *Authority and Estrangement: An Essay on Self-Knowledge*, Princeton, Princeton University Press.

Moran, R. (2003), “Responses to O’Brien and Shoemaker”, *European Journal of Philosophy* 11, pp. 401-419.

Moya, C. (2007), “Moran on Self-Knowledge, Agency and Responsibility”, *Crítica* 38, pp. 3-20.

Shoemaker, S. (2003), “Moran on Self-Knowledge”, *European Journal of Philosophy* 11, pp. 391-401.

## Directness or saliency? On the semantics of indexicals

Aldo Frigerio  
Università Cattolica di Milano  
Department of Philosophy  
Largo Gemelli 1  
20123 Milano - Italy

As is well known, Kaplan has claimed that indexicals are direct reference devices. However he does not think that the directness of a designator implies that there is no semantic mediator between the designator and the referent. On the contrary, the character of an indexical is part of its semantics, despite being a function from the context to the referent and, therefore, acting as a mediator between the designator and its referent. As Genoveva Marti has pointed out, in Kaplan view, to say that a reference device is direct does not mean that it is Millian (i.e. that its meaning is nothing but its referent), but it means that only the referent enters into the truth conditions of the sentences containing the device. Even if the character of an indexical serves as a means to pick up the referent and it is part of the meaning of the sentences in which the indexical occurs, the character does not enter into the truth conditions of those sentences.

Kaplan argues that the sentences containing indexicals express Russellian singular propositions on the basis of the behavior of indexicals in presence of modal, temporal and spatial operators. In these cases, the character of the indexical is never evaluated in a world, time and place different from those of the speaker's. These operators are not able to turn the evaluation index of the character the way they do with the evaluation index of the descriptive content of a definite description. For this reason, while definite descriptions may have wide or narrow scope with respect to these operators, indexicals always take the largest scope.

We can give two interpretations of this behavior of indexicals when co-occurring with some operator: (i) they express a Russellian proposition: the operator cannot handle their character because the character does not enter into the proposition and only the referent does. (ii) the character enters into the proposition, which is a descriptive proposition, but for some reason the operator always takes narrow scope with respect to the descriptive content of the indexical. Kaplan refuses (ii) and accepts (i) because he sees no reason why a descriptive content should always take the widest scope once loaded in a proposition. Yet I would like to explore a different view: (ii) is true and actually there is a reason why indexicals always take the largest scope. According to this view the semantic peculiarity of indexicals with respect to definite descriptions does not consist in the directness of their reference but in what I call saliency. In contrast to definite descriptions, indexicals must refer only to something salient in the context of utterance, i.e. to something the speakers are paying attention to during their conversation. I believe that there are two kinds of saliency, which I call natural saliency

and artificial saliency. An object is naturally salient in a context when for some reason the speakers are led to pay attention to this object. A very loud noise, a bright light, a shouting man, just to name a few instances, are naturally salient objects. Referring to objects of this kind does not usually require any kind of deictic gesture. If a loud noise suddenly arises, a speaker can say: “What is this noise?” without any deictic gesture (cfr. also Kaplan’s example of the man running away and of the speaker who says: “Stop that man!” without any gesture). On the other hand, artificial saliency is due to deictic gestures, which I interpret as means for making salient what is not salient (enough) to be the referent of an indexical. The speaker draws her addressees’ attention on something by making a deictic gesture which in turn makes it salient.

How can saliency explain the behavior of indexicals in presence of modal and temporal operators? In order to be salient for both the speaker and the addressee, an object must be present in the utterance context. Something is salient only when both the speaker and the addressee are paying attention to it and are mutually aware that the other is doing the same. This condition is satisfied only when there is an object that stands out for some reason. But an object can both stand out and draw the attention of all the speakers taking part to a conversation only if it’s part of the utterance context. Objects from other worlds, times and places cannot draw all the speakers’ attention at the same time. But something that is part of the utterance context should be recovered in the world, time and place of the utterance. Since the referents of indexicals must be part of the context, these devices of reference cannot take narrow scope. This suffices to account for the behavior of indexicals in presence of operators.

In my opinion, there are many advantages to the reference to salient objects being the main feature of indexicals:

- 1) It accounts for the fact that we prefer to use definite descriptions to refer to non-salient objects: \*“this first man who went to the Moon” vs. “the first man who went to the Moon”;
- 2) It accounts for the fact that we prefer to use indexicals to refer to salient objects: “stop that man” vs. ?“stop the man”;
- 3) Indexicals are usually considered ambiguous between an anaphoric (endophoric) use and an exophoric use. But in the view defended here it is not necessary to postulate this ambiguity: indexicals always refer to salient things. But a thing can be salient either because it is present in an extralinguistic context or because it is the topic of the speakers’ conversation, namely it is present and salient in a linguistic context. Indexicals refer to both kinds of salient things. Notice that an object which is the topic of a conversation needs not to be in the utterance (extralinguistic) context; hence anaphoric indexicals need not to take wide scope with respect to modal and temporal operators as exophoric indexicals do.

## Conditional Requirements of Rationality

Antonio Gaitán Torres  
Universidad de Granada  
[agaitan@ugr.es](mailto:agaitan@ugr.es)

Conditional requirements of rationality concern the attitudes that an agent is required to form (or to drop) only in virtue of a given subset of her attitudes. As example:

(*Instrumental Principle*) Rationality requires intending F if you intend E and you also believe that if E then F and that F will only be the case if you intend so

There is a wide agreement on the availability of conditional requirements of rationality in the sense noted above. Such agreement vanishes, however, when we look for an explanation of their peculiar *force* or *normativity*. The normativity of conditional requirements of rationality will be the topic of this talk.

An obvious route to explain the force of conditional requirements would assume that every conditional requirement is normative in every occasion. Otherwise:

(**NR**) Necessary, if you are rationally required to F by a given requirement of rationality, then you have a conclusive reason to F so that you ought to F

Despite its intuitive plausibility, NR faces a problem. The problem is better stated by considering an example. Let us suppose that you intend to kill Juan, believing also that poisoning his drink would be the best way to do so. We can assume a lot of things about this case. Some of them are normative. We can assume something about the merit of your attitudes and surely something about the kind of person you are. But crucially we can also assume something normative about you if we ignore the merit of your attitudes or the kind of person you are. We can assume, in a nutshell, that a rational requirement in the above sense can be applied to you. The problem is that if the requirement is a Narrow-scope requirement

(**NS**)  $I_p \wedge B(p \rightarrow q) \rightarrow RRIq$

and we assume NR

(**NS**) + (**NR**)  $I_p \wedge B(p \rightarrow q) \rightarrow OIq$

then all the above talk about ‘pressure’ and ‘requires’ would simply mean that from your current intentions and beliefs it follows an unconditional obligation in favour of poisoning Juan’s drink. It is so because from NS + NR we can *detach* a normative consequent once we assume that the antecedent is the case – in the present case your intention and instrumental belief about Juan. It seems clear to me, however, that we don’t want to explain the pressure you are under in this case by endorsing the advice offered by NS + NR. Even if your attitudes demand by themselves so to speak something of you, it cannot be your poisoning Juan’s drink. You ought not to poison Juan’s drink and because of that NS + NR cannot offer the right way to unfold the force of conditional requirements of rationality.

An alternative to NS + NR would claim that the requirement you are under in the above case is a Wide-scope requirement

(WS) RR  $(Ip \wedge B (p \rightarrow q) \rightarrow Iq)$

According to this option [Broome, J. 2007], when we conjoin WS and NR

(WS) + (NR) O  $(Ip \wedge B (p \rightarrow q) \rightarrow Iq)$

it does not follow that you ought to poison Juan's drink. That would be invalid *detachment* (because the scope of 'ought' in WS + NR is the entire conditional). What you ought to do here according to WS + NR is either to intend to poison Juan's drink or to abandon *any* of your grounding attitudes (intending to kill Juan or believing that if you intend to kill Juan, then you have to poison his drink). This logical manoeuvre fits much better with our intuitions; after all, we are disposed to recommend you to drop your intention to kill Juan in this case and this precise advice can be derived from WS + NR. The conclusion is that if we want to explain the force of conditional requirements of rationality in terms of NR we should prefer WS over NS as the natural way to model the logical form of conditional requirements of rationality.

The conjunction of NR and WS has been recently attacked on two different fronts. It has been argued, first, that the *logical symmetry* inside WS cannot accommodate the *rational asymmetry* inside conditional requirements [Schroeder, M. 2007]. And it has been stressed, second, that WS + NR cannot offer an explanation of why conditional requirements are normative *in every occasion* [Kolodny, N. 2008]

In my talk I will defend WS + NR against both objections.

## REFERENCES

- Broome, J. 2007. "Wide or Narrow Scope?" Mind, 107, pp. 359-370
- Kolodny, N. 2008. "Why Be Disposed to be Coherent?" Ethics, Vol. 118, July, pp. 437-463
- Schroeder, M. 2009. "Means-end Coherence, stringency, and subjective reasons", Philosophical Studies, 143, pp. 223-248

## Una concepción (neo)expresivista del autoconocimiento

Ángel García Rodríguez

Universidad de Murcia

agarcia@um.es

El objetivo de esta comunicación es discutir la concepción expresivista del autoconocimiento asociada con la filosofía de Wittgenstein. En concreto, a la luz de algunos trabajos recientes sobre el expresivismo, y su relación con Wittgenstein, el objetivo principal es determinar si una concepción expresivista del autoconocimiento ha de renunciar a la noción de verdad; o si por el contrario, se puede encontrar algún vínculo entre las nociones de expresión y verdad que pueda arrojar luz sobre el problema del autoconocimiento.

En *El cuaderno azul*, Wittgenstein defiende que las emisiones (al menos típicas) de “yo tengo dolor de muelas” no tienen la estructura lógica sujeto-predicado, ni por tanto la función comunicativa de aseverar algo de alguien, sino más bien la de expresar la condición mental del sujeto. Llamemos “expresivismo” a esta concepción del significado de las autoadscripciones de estados mentales típicos en primera persona del presente (o simplemente, autoadscripciones). En la medida en que la noción de verdad está estrechamente vinculada con la de aseveración, se asume a veces que la concepción expresivista del significado de las autoadscripciones implica que, *qua* expresiones, aquéllas carecen de valor de verdad. De aquí se concluye que, si (como es generalmente aceptado) el conocimiento implica verdad, entonces las autoadscripciones no son afirmaciones genuinas de autoconocimiento. En este sentido, el expresivismo es una alternativa a otras concepciones (contemporáneas y clásicas) del autoconocimiento, que no cuestionan que las autoadscripciones tengan valor de verdad.

La idea de que, *qua* expresiones, las autoadscripciones carecen de valor de verdad descansa sobre el modelo proporcionado por las expresiones naturales o no lingüísticas. Sin embargo, recientemente se ha negado que la concepción de las autoadscripciones como expresiones implique que carezcan de valor de verdad. Así, en *La expresión y lo interno*, Finkelstein defiende que, como sugiere el propio Wittgenstein, aunque las autoadscripciones se parecen a las expresiones naturales, “no son *exactamente* como las expresiones faciales o corporales: hay una diferencia crucial, que consiste en que tienen una dimensión asertiva y pueden ser verdaderas o falsas” (224). La clave, para Finkelstein, es que aseverar y expresar no se excluyen mutuamente, con lo que el carácter expresivo de las autoadscripciones no es un obstáculo para que tengan valor de verdad. Sin embargo, Finkelstein no termina de explicar en qué sentido la noción misma de expresión se relaciona con la de verdad.

De manera parecida, en *Speaking my mind*, Bar-On rechaza el expresivismo simple, según el cual las autoadscripciones son expresiones en el mismo sentido que las expresiones naturales, en favor de una concepción neoexpresivista que distingue entre el acto realizado y el producto de ese acto. Así, según Bar-On, las autoadscripciones son similares a, y continuas con, las expresiones naturales *qua* actos; pero se diferencian en que sólo en el caso de las autoadscripciones el producto de ese acto es lingüístico, y por lo tanto susceptible de verdad y falsedad. Sin embargo, con esta nítida separación entre acto y producto, se separa el componente de expresión de las autoadscripciones del

componente de verdad; y en definitiva tampoco se consigue explicar en qué sentido el carácter expresivo de las autoadscripciones se relaciona con la noción de verdad.

Tanto Finkelstein como Bar-On quieren marcar las diferencias entre el significado de las expresiones naturales y el de las autoadscripciones lingüísticas, pues éstas tienen condiciones de verdad de las que aquéllas carecen. Quizás podría marcarse esta diferencia diciendo que las expresiones naturales pueden ser sinceras o no, mientras que las autoadscripciones pueden ser verdaderas o falsas. Aunque tal vez sería más apropiado marcar la diferencia diciendo que sólo las autoadscripciones pueden ser verdaderas o falsas, dado su carácter lingüístico. Pues parece claro que, por su carácter expresivo, las autoadscripciones, lo mismo que las expresiones naturales, pueden ser sinceras o no. El resultado de esto es que el significado de las autoadscripciones vendría dado por una suerte de combinación de dos tipos de condiciones distintas: a saber, condiciones de sinceridad más condiciones de verdad.

Una alternativa a esta forma de entender las propuestas (neo)expresivistas contemporáneas sería establecer un vínculo más estrecho entre las nociones de autoadscripción y verdad. Para ello, no habría que renunciar al vínculo señalado anteriormente entre expresión y sinceridad, sino mostrar que hay un vínculo entre sinceridad y verdad. Así, partiendo de la tesis siguiente:

(i) Una autoadscripción es expresiva si y sólo si es sincera,

podría explicarse el vínculo entre expresión y verdad de la manera siguiente:

(ii) Una autoadscripción es sincera si y sólo si se da el estado de cosas (mental) correspondiente.

Una manera de argumentar a favor de (ii) es por contraposición con una concepción alternativa de la sinceridad de una autoadscripción, según la cual:

(iii) Una autoadscripción es sincera si y sólo si el sujeto *cree* que se da el estado de cosas (mental) correspondiente.

El objetivo de esta comunicación es defender (ii). Para ello se argumenta que (iii) no da condiciones necesarias ni suficientes a favor de la sinceridad de una autoadscripción. Así, una autoadscripción puede ser sincera y por tanto expresiva, a pesar de que el sujeto en cuestión no crea que se da el estado de cosas (mental) correspondiente (quizás el sujeto carezca de creencias al respecto); con lo que la creencia del sujeto no es una condición necesaria de la sinceridad de una autoadscripción. Pero tampoco es una condición suficiente, pues resulta contraintuitivo pensar que la existencia de la creencia del sujeto basta para que una autoadscripción sea expresiva. Más bien, parecería que es crucial que, *además*, se dé el estado de cosas (mental) correspondiente, el estado expresado. Pero si, como se ha señalado antes, la creencia del sujeto respecto a si se da el estado mental correspondiente no es condición necesaria para la sinceridad, y por tanto para el carácter expresivo, de una autoadscripción, lo que resulta es que la existencia del estado mental correspondiente es la condición suficiente que buscábamos. Con lo que se habría defendido la opción (ii).

## Sobre causalidad y la dirección medios-fines en la acción

M.J. García-Encinas  
Dpto Filosofía I, Facultad de Psicología  
Campus Cartuja, Universidad de Granada

El objeto de esta comunicación es desentrañar algunos aspectos fundamentales de la relación entre los conceptos de acción y causalidad. En líneas generales, quiero defender que aunque acción y causalidad pertenecen a reinos ontológicos diferentes, es posible establecer importantes relaciones lógicas entre ellos. En particular, la relación medios-fines que es esencial a la acción puede ser la relación causal, si medios y fines se individuán en términos de propiedades (o hechos) en vez de en términos de sucesos. Así, aunque las acciones no son propiamente causas y efectos, los fines y medios que las conforman sí pueden entrar en relaciones causales. La causalidad proveería, por tanto, la direccionalidad de la acción, siendo los medios las causas de los fines.

Hay dos grandes tipos de teorías que relacionan acción y causalidad, y lo hacen inversamente: las *teorías causales de la acción*, y las *teorías agenciales de la causalidad*. A las primeras las distingue el mantener la tesis:

(CA) Las acciones son causadas y causas.

Puesto que según (CA) las acciones forman parte de la ontología causal, en términos explicativos y ontológicos la relación causal es primera a la acción. No hay acción sin causalidad. (CA) ha sido especialmente defendida por Davidson (1963), pero muchos otros filósofos la comparten: Richard Taylor (1973), O'Shaughnessy (1973), Hornsby (1980), Searle (2001), o Richard Wright (2007). Por otro lado, a las teorías agenciales de la causalidad les une la defensa de:

(AC) La causalidad es (se deriva de la) agencialidad.

Según las teorías agenciales, el concepto de causalidad es secundario y dependiente de nuestra condición de agentes en el mundo. Enmarcadas en una tradición de corte psicologista-humeano, las diversas teorías agenciales de la causalidad mantienen que, en sentido estricto, no hay causalidad en el mundo. Actuamos y nos percibimos como agentes, y así proyectamos nuestra capacidad agente sobre el mundo, interpretando la realidad como si efectivamente los hechos se sucediesen unos como consecuencia de otros. Esta es la posición que tradicionalmente representan Gasking (1956) y Ayer (1956) y, contemporáneamente, Huw Price (1996).

Aquí quiero posicionarme en contra de ambas tesis. Contra (CA) mantengo que las acciones no conforman la ontología de la relación causal porque los sucesos no conforman la ontología causal; son más bien propiedades (o hechos) los que entran en relaciones causales. Junto a filósofos de corte más Wittgensteiniano, igualmente detractores de las teorías causales de la acción, como Malcolm (1968), von Wright (1971) o Charles Taylor (1979), defenderé que acción y causalidad pertenecen a reinos ontológicos distintos. Las acciones no causan ni son causadas, no se explican causalmente o en términos nomológicos, sino que se comprenden. Pero (AC) también yerra; y lo hace en la medida en que asume que la causalidad es un concepto derivado. Contrariamente a los teóricos agenciales de la causalidad, defiende que hay relaciones causales en el mundo; hay hechos causales básicos y primitivos. La relación causal no se reduce ni analiza en términos de

otras relaciones o conceptos, menos aún es un pseudo-producto de nuestra psicología. No obstante, hay algunas intuiciones certeras en (AC). En particular, una idea que es central a las teorías agenciales de la causalidad y según la cuál la asimetría causal se aprehende en la dirección medios-fines de la acción. Causas y efectos no pueden intercambiar sus roles causales: si  $c$  es causa de  $e$ , entonces  $e$  no es causa de  $c$ . Esta dirección interna en la causalidad sólo podemos aprehenderla en la medida en que podemos interactuar con un mundo primeramente causal. Es decir, si  $c$  es causa de  $e$ , entonces llevar a cabo  $c$  es una buena estrategia medios-fines para conseguir  $e$ . Para defender estas ideas voy a rescatar de lo que considero un injusto olvido algunas de las ideas de von Wright (1971, 1973, 1974, 1976).

En *primer lugar*, presentaré algunas consideraciones sobre la acción y defenderé una concepción mínima de la misma, según la cuál una acción es un suceso del que un agente es responsable. Es decir, una acción es un tipo de suceso; un suceso atribuible o imputable a un agente. Dado el rechazo de las teorías causales de la acción, la idea de responsabilidad agente no puede ser causal, y también se aleja de toda concepción de corte davidsoniano que liga la acción de forma esencial a la intencionalidad o la racionalidad. Por el contrario, la idea de responsabilidad agente que manejaré es social y va de la mano de las nociones de responsabilidad moral y de revocabilidad de Hart (1948/9) y Feinberg (1965). En *segundo lugar*, recuperaré algunas tesis fundamentales de la teoría de la acción y de la causalidad de von Wright; en concreto, la lógica que él defiende para la relación entre acción y causalidad, en el intento de poner sobre la mesa los ejes que gobiernan el encuentro entre ambos reinos. En *tercer lugar*, defenderé que estas ideas permiten mantener la tesis de que la asimetría causal se refleja en la dirección medios-fines de la acción. Si  $c$  es causa de  $e$ , entonces llevar a cabo  $c$  es una buena estrategia medios-fines para conseguir  $e$ , donde  $c$  y  $e$  son resultados (medios) y consecuencias (fines) de acciones. Más aún, defenderé que esta direccionalidad es compatible con la posibilidad del establecimiento de relaciones causales temporalmente invertidas, incluso cuando éstas puedan establecerse en el seno de la acción. Es quizás mediante nuestra condición de agentes que tenemos acceso epistemológico al establecimiento de relaciones causales en el mundo. Sin embargo, la causalidad no es, ni se deriva de, nuestra condición de seres agenciales.

## What is said in irony?

Joana Garmendia  
ILCLI (University of the Basque Country)  
Villa Asunción  
Avda. José Elosegui, 275,  
20015 Donostia-San Sebastián  
[joana.garmendia@gmail.com](mailto:joana.garmendia@gmail.com)

Explaining ironic communication from a pragmatic standpoint has proved to be a challenge as interesting as it has been tricky. Some leading theories got off to a promising start in this endeavor, but they ended up making some rather forced maneuvers when counterexamples were presented one after the other.

The traditional pragmatic accounts avoid using the term “saying” when explaining irony. This is not a mere coincidence –behind that avoidance strategy a fundamental characteristic of irony lies hidden: ironic utterances cannot be easily reported with common verbs such as “saying”, “asserting” or the like. That’s what we call irony’s “*what is said* issue”: it really seems that nothing is *said* in irony. That’s the challenge every pragmatist must face when entering the analysis of ironic communication.

In this paper, I shall mention the main alternatives offered up to now to solve this issue: Grice (1967/1989) states that the ironic speaker *makes as if to say*; Sperber & Wilson (1981) claim that she makes *an echoic mention*; and Clark & Gerrig (1984) hold that she *pretends to say*. If considered separately, each of these approaches has some major limitation, which makes us reject it as a valid option for facing the issue. But, if we look at them as a whole, we shall realize that they take a very similar approach to tackling the issue we are now concerned with: instead of clearing up the concept of *saying* in ironic utterances, they merely get around the problem, and offer a substitute for “ironically saying”.

I shall then propose an alternative theory to explain irony, following the pathway opened up by Korta & Perry (2006a, 2006b, 2007a, 2007b, 2007c) in their *Critical Pragmatics*. In fact, we must realize that the problem of *what is said* does not belong exclusively to ironic utterances, but comes from the existing confusion within general pragmatics –as a consequence of the issue regarding the Pragmatics-Semantics interface. We shall see that the critical pragmatic approach clears up this general mess, and so can help us solve irony’s specific problem.

Critical Pragmatics distinguishes several contents within an utterance, and multiple layers within the speaker’s communicative plan. Therefore, a more detailed analysis of utterances is provided. This makes a notable advantage when explaining any case of

communication, even the most paradigmatic cases of speech (where the speaker is usually considered to be sincere, literal and error-free), but this advantage is made more noticeable when facing cases that depart from the paradigmatic ones: Critical Pragmatics offers the perfect basis for explaining cases where the speaker has been insincere, cases where she has made a mistake or where she has communicated a non-literal meaning.

It's not surprising thus that analyzing ironic communication becomes an easier task once we adopt the basic claims of Critical Pragmatics. As soon as we acknowledge that multiple contents can be distinguished within an ironic utterance, the lack of an identifiable "what is said" doesn't seem such a theoretical burden. I will show that this is so by proposing a basic account for ironic utterances: it's named the *Asif*-Theory, it's basically Gricean, and it can explain how an ironic speaker can successfully communicate even if she doesn't *say* any of the contents of her utterance. We will see that the speaker, when being ironic, *makes as if to say* a content instead, and thereby implicates a different content, the ironic one.

Summing up, I shall maintain that adopting Critical Pragmatics as our general pragmatic basis allows us to dismantle the "*what is said* issue", while offering a suitable explanation for ironic communication.

## References

- Clark, Herbert H. & Richard J. Gerrig (1984), "On the Pretense Theory of Irony". *Journal of Experimental Psychology: General* **113** (1): 121-126.
- Grice, Paul (1967), "Logic and Conversation". In D. Davidson & G. Harman (eds.) (1975), *The Logic of Grammar*. Encino: Dickenson, pp. 64-75. Also published in P. Cole & J.L. Morgan (eds.) (1975), *Syntax and Semantics 3: Speech Acts*. New York: Academic Press, pp. 41-58. Reprinted in P. Grice (1989), pp. 22-40.
- Grice, Paul (1989), *Studies in the way of words*. Cambridge (MA): Harvard University Press.
- Korta, Kepa & John Perry (2006a), "Three demonstrations and a funeral". *Mind and Language* **21** (2): 166-186.
- Korta, Kepa & John Perry (2006b), "Varieties of Minimalist Semantics". *Philosophy and Phenomenological Research* **LXXIII** (2): 451-459.
- Korta, Kepa & John Perry (2007a), "Radical minimalism, moderate contextualism". In G. Preyer & G. Peter (eds.), *Context-Sensitivity and Semantic Minimalism. Essays on Semantics and Pragmatics*. Oxford: Oxford University Press, pp. 94-111.
- Korta, Kepa & John Perry (2007b), "How to Say Things with Words". In Savas L. Tsohatzidis (ed.), *John Searle's Philosophy of Language: Force, Meaning, and Thought*. Cambridge: Cambridge University Press, pp. 169-189.
- Korta, Kepa & John Perry (2007c), "The pragmatic circle". *Synthese*. Online first, DOI: 10.1007/s11229-007-9188-3).
- Sperber, Dan & Deirdre Wilson (1981), "Irony and the use-mention distinction". In P. Cole (ed.), *Radical Pragmatics*. New York: Academic Press, pp. 295-318.

## The function of global scenarios in Wittgenstein's *On Certainty*

Modesto M. Gómez Alonso

Universidad Pontificia de Salamanca

One particular version of skepticism, the so-called Cartesian skepticism (a variety of skepticism which deploys such scenarios as those proposed by Descartes in the First Meditation), was under constant attack during the last century. Austin, Bouwsma, Ryle or, in more recent times, Michael Williams, Barry Stroud and Stanley Cavell, have been so effective that almost nobody holds this position nowadays. What is problematic with Cartesian skepticism?

Mainly two things: that the skeptical scenarios really are reasons for doubting our *ordinary beliefs*, that is to say, that they are used by radical philosophers in order to challenge our common beliefs of everyday life (“Here is a hand”, “I’m a human being”, “Physical Objects continue to exist when unperceived”...), something which cannot be done without endangering the very conditions of a reasonable doubt; and the fact that traditional skeptical arguments are (allegedly) either committed to obscure, bizarre and meaningless philosophical doctrines and contentious entities (for instance, sense data and internalist conceptions of meaning) or they are what Michael Williams called “unnatural doubts”, those resulting from the employment of language under conditions stipulated arbitrarily by the philosopher, such that the language so employed has ceased to have any meaning. In other words, owing to the *global scenarios* it has to use in order to raise an *unmitigated skepticism*, Cartesian skepticism is defective. The dream argument is a good example of global scenario. If this argument provides a reason for doubt the present experience, then, because if I may be dreaming now I may be dreaming at any time, it provides also a reason to doubt whatever experience we appeal to in order to rule out that possibility (I could be dreaming of shaking my head or pinching my face as means to settle the question whether I’m fast asleep or awake). What were the consequences of these criticisms? Obviously, Cartesian skepticism was rejected. But, because it was considered that Cartesian skepticism was the *only possible sort* of general and philosophical skepticism, this rejection amounted to a *redirection* of epistemology: skepticism was ruled out from debates; among epistemologists it was a common assumption that we *do possess* knowledge, and thereby that the task of a theory of empirical justification was *not* to show *that* knowledge is possible, but *how* it is possible.

Fifteen years ago Robert Fogelin burst this fixed situation, refusing to equate general skepticism with Cartesian skepticism. Like Sextus Empiricus in the ancient world he championed Pyrrhonian skepticism in the contemporary world. His goal was to reintroduce the question about if *knowledge in general* is possible in epistemological debates, task which required a sort of skepticism (at least) as radical as Cartesian skepticism but without the caveats and commitments which turned this position (because it might be possible to bring forth arguments showing that classical skeptical scenarios are conceptually incoherent) into a vulnerable one. In other words, Pyrrhonian skepticism *had to be* immune against the charge of *meaninglessness* without losing

momentum. It is obvious that this Pyrrhonian skepticism doesn't achieve its aim: to get as robust a skeptical challenge as one could like only through checkable but unchecked defeaters, that is to say, to raise eliminable doubts equally devastating than uneliminable global doubts. In order to demonstrate this point it is enough to remember why Descartes *had to use* the dream hypothesis to extending skepticism: because there are "many other beliefs about which doubt is quite impossible...for instance, that I'm here, sitting by the fire, wearing a winter dressing-gown, holding this piece of paper in my hands." It's quite evident that if I could doubt that this is my hand or that I'm writing this abstract, my grounds to doubting *wouldn't be empirical conditions* which I have to eliminate in order to know, but a global scenario which, though maybe I could eliminate, I couldn't rule out appealing to data within its scope. In other words: there are *nuclear beliefs* beyond the scope of Pyrrhonian arguments, and thus, either this sort of skepticism is not radical enough or it must appeal to Cartesian scenarios in order to be so.

The previous fact is closely related to the *extraordinary interest in global scenarios* which Wittgenstein showed throughout the remarks which were published by his trustees in 1969 under the title of *On Certainty*. Why this infatuation? Because, unlike unchecked defeaters, which question a claim because *it is not grounded enough* but which don't question *what counts as evidence to solving the question we accidentally cannot solve*, namely, which don't raise doubts on the meaning of "to be grounded"; global scenarios defeat our claims *defeating whatever evidence we could give* to defend them in the arena of reflection, and thereby *defeating the very notion of evidence, which, paradoxically, lacks evidence*. What they question is our cognitive system *as a whole*. Which means that the Cartesian skeptic doesn't say that we don't have such and such experiences which *would remove* our doubts, but that there is *no experience* to appeal to in order to rule them out. Ironically, it is this *radical feature* what, enticing Wittgenstein, makes of *On Certainty* the only significant work in Twenty Century epistemology (previous to the strand known as *virtue epistemology*) which *vindicates global scenarios*. In short, the most important achievement of Wittgenstein's last remarks was the reintroduction and *partial rehabilitation* of Cartesian skepticism in contemporary philosophy. What is the meaning of "partial rehabilitation of Cartesian skepticism"? This question might be glossed as: What are the *roles* played by global scenarios in *On Certainty*? To identify these roles means to define the main targets Wittgenstein is aiming at in *On Certainty*, which, in my opinion, are *three*: (i) *rejection of foundationalism and evidentialism*; (ii) *reconstruction of the skeptical problem at a deep enough reflective level* (which entails that global scenarios are *context-dependent*); and, (iii) *vindication of our reliance in hinge-propositions when a doubt over them is rational*.

I'll divide this paper in three parts. First, I'll give an *outline* of the fate of skepticism in contemporary analytic philosophy, thus providing the necessary *background* in order to understand the epistemological conditions Wittgenstein and his interpreters are responding to. In second place, I'll face four possible readings of *On Certainty* (epistemological naturalism, transcendental interpretation, therapeutic and Neo-Pyrrhonian readings), all of them implicit in the previous preliminary account. Finally, I'll describe what I consider the *very heart* of Wittgenstein's remarks, a novel and extraordinary version both of *radical skepticism* and *reliabilism* which I'll label with the barbarous word "*inverted reliabilism*".

## **Intersubjectivity and Empathy**

Antoni Gomila

Dep. Psicologia - Universitat de les Illes Balears

antonigomila@gmail.com

### **Abstract**

In this presentation I will take issue with the successful standard functional interpretation of "mirror neurons" as the carriers of intersubjectivity. This interpretation, dependent upon the Simulation Theory of intentional attribution, understands the double duty of such neurons as the implementation of the mechanism that allows to project onto others one's own intentional states, and according to more recent proposals, one's own emotional states, thus turning empathy as a prominent area of research. Against this view, I will contend: a) that Simulation Theory doesn't provide an account of intentional imitation or empathy in general; b) in particular, that by "empathy" a collection of diverse phenomena is covered, and thus, not distinguished; c) that some of these phenomena cannot be accounted for in simulationist or theoretical terms, but call for a "second-personal" view of intersubjectivity as a ground level of interaction; and finally, d) that the fact that a neuronal structure is involved in several mental functions cannot be seen as something peculiar to the "mirror system", but needs a principled explanation given that it is the rule of brain organization.

### **Proposal**

The discovery of "mirror neurons" in the 90s has brought about an interest in intersubjectivity and the mechanisms that make intentional interaction possible. These neurons were localized in the premotor cortex of the macaque, and are characterized by the fact that fire both when the subject perceives a certain kind of intentional action (grasping or pointing), and also when the subject itself performs that very action (Rizzolatti & Sinigaglia, 2006). Although the existence of "mirror neurons" has not been demonstrated in humans yet, because the procedure would require "in vivo" single neuron recording, neuroimage evidence supports the idea of a "mirror system" also in the human premotor cortex, with similar double functionality.

From very early on, this neurobiological finding has been interpreted as providing support for the Simulation Theory of mental attribution (Gallese & Goldman, 1998; Gallese, 2001; Hurley, 2008). According to this interpretation -which has become standard to the point of being presented as "common knowledge" (Iaconi, 2008)-, these double duty neural structures provide the mechanical implementation of the simulational process taken to be the basic, central, procedure for mental attribution. In this way, Simulation Theory has renewed and updated the classical analogical theory of intersubjectivity.

More recently, this view has been extended to emotional interaction: one grasps what another feels by being activated in the same way as when one feels like that and projecting that feeling to the other (Goldman & Sripada, 2005). While there is discussion whether this "emotional mirror system" completely overlaps with the

"intentional mirror system" or involves different neural structures (Decety, 2010), this move has renewed the interest in empathy as a basic form of intersubjectivity (Singer, 2006; Adolphs, 2002).

In the past, I have opposed Simulation Theory as a proper way to account for intersubjectivity (as well as Theory Theory), and suggested a different tack: a second-personal approach to intentional interaction. In this paper, I will develop this alternative by:

- 1) showing that Simulation Theory falls prey to the classical objections against Analogy as a solution to the classical "knowledge of other minds" problem (Zahavi, 2008; Gallagher, 2008);

- 2) showing that empathy covers phenomena that are not accountable in terms of "same feeling shared by attributor and attributee" and projection: sympathy, empathic projection, identification, not even emotional contagion (no projection involved); that case is rather the exception;

- 3) arguing that the second person perspective (Reddy, 2008) is the right framework to account for the ontogenetic data about intentional development; according to this view, intentional interaction starts from expressive interaction, and gives rise to semiotic and linguistic interaction; propositional intentional attribution is the culmination of the process, not the starting point;

- 4) pointing out that multifunctionality is the rule in brain organization, not something extraordinary related to intentional interaction.

## Is “pre-reflective self consciousness” a solution for the problem of the self?

Jorge Gonçalves  
Rua das Adelas, 24 r/c 1200-008 Lisboa  
jorgalvesenator@gmail.com

In this paper I will present an objection to the solution to the problem of the self proposed by Dan Zahavi in "Is the self a social construct?" *Inquiry* 52/6, 2009, 551-573, and other texts. This objection is based on developmental psychology of the human individual.

I phenomenally appear to myself as an entity that remains in time and differentiates itself from other selves. The existence of such an entity has been questioned by several philosophers: Hume, Nietzsche and more recently Metzinger *The Ego Tunnel: The Science of the Mind and the Myth of the Self* (2009).

This last author explicitly denies the existence of the self: "the conscious self is the illusion which is in one's illusion" and "no such things as selves exist in the world: Nobody ever was or had a self". We have the illusion of being a self or of having a self because we have no knowledge of the sub-personal representational mechanisms and neural mechanisms that create the process. It is a hallucination created by a system that does not recognize it as such:

“ No such things as selves exist in the world. For all scientific and philosophical purposes, the notion of a self – as a theoretical entity – can be safely eliminated. What we have been calling "the" self in the past is not a substance, an unchangeable essence or a thing (i.e., an "individual" in the sense of philosophical metaphysics), but a very special kind of representational content: The content of a self-model that cannot be recognized as a model by the system using it. ” Metzinger “Précis: Being No One” *Psyche* 11 (5), June 2005.

Dan Zahavi proposes a solution for the problem of the self. Firstly, he distinguishes the notion of narrative self from the notion of core or minimal self. Gallagher defined these two kinds of self in a succinct manner: “Narrative self: a more or less coherent self (or self-image) that is constituted with a past and a future in the various stories that we and others tell about ourselves” and “Minimal self: Phenomenologically, that is, in terms of how we experience it, a consciousness of one self as an immediate subject of experience, unextended in time”.

Zahavi’s idea, supported in continental phenomenology, is that phenomenal consciousness and the “what is it like” character of experience always presupposes a pre-reflective self-consciousness.

“The crucial idea propounded by all of these phenomenologists is that an understanding of what it means to be a self calls for an examination of the structure of experience, and vice versa. Thus, the self is not something that stands opposed to the stream of consciousness, but is rather immersed in conscious life; it is an integral part of its structure. More precisely, the (minimal or core) self is claimed to possess experiential reality, it is taken to be closely linked to the first-person perspective, and is in fact identified with the first-personal *givenness* of the experiential phenomena. This first-personal givenness of experiential phenomena is not something quite incidental to their being, a mere varnish that the experiences could lack without ceasing to be

experiences. On the contrary, this first-personal givenness makes the experiences *subjective*.” Zahavi “Self and other” (2007).

A consequence of Zahavi’s theory is that all conscious experiences will have to have this property of “mineness”. The experience will be felt by me as being different from others. Of course, I don’t need to think of myself as an entity or a thing in analogy with material things to have an experience. However, I need to feel myself minimally as a separate “I” of the other “Is”. The experiences have the character of the first person “givenness”. It is not possible to talk about experience without someone having those experiences.

If this is true then any creature that has phenomenal consciousness will have to have that pre-reflective self-consciousness. The same will be true for the human being in his early stages of life.

Now, I believe that the scenery according to which babies begin their lives without any sensation of being subjects but still having psychological experiences is perfectly plausible. Life in these early stages seems to me to be a “chaos” of sensations (from our point of view) without unity and the feeling of being a separate subject from others. In this case, we cannot consider that there is any self-consciousness even in the sense of a pre-reflective self-consciousness.

Which are the proofs that can confirm this empirical conjecture? As this belongs to the first person epistemological domain, the objective proofs will hardly be completely conclusive since we cannot prove what the phenomenal states of babies are in fact. We can observe the behaviour of the babies but not observe what are their mental experiences. However, some authors from several theoretical backgrounds such as Alison Gopnik (*The Philosophical Baby*, 2009) in the area of analytical philosophy or some psychoanalysts hold, based on experimental data, that in the initial stages of life there is not trace of a feeling of a separate I.

If we admit that initially in life there is no pre-reflective self-consciousness, we have to admit that the feeling of self is a process formed by maturation and interaction with other people. The self, even the core or minimal self, is therefore a mental construction that forms itself from pre-existent materials, that is, from phenomenal pre-subjective states. This conclusion is speculative but I believe that it’s compatible with the neurological and the behavioural facts.

Can we then affirm that the self is an illusion? I think not. It is more correct to speak about “mental construction” because we cannot assume a priori that the ontology of mind is the ontology of illusion and that the only reality is the brain.

## The Paradox of Pictorial Presence

Geert Gooskens

University of Antwerp

Department of Philosophy

Grote Kauwenberg 18 (office: D-412)

2000 Antwerp, Belgium

[geert.gooskens@ua.ac.be](mailto:geert.gooskens@ua.ac.be)

+32 32654325

<http://www.ua.ac.be/main.aspx?c=geert.gooskens>

What exactly is the object of our experience when we look at a picture? This question is tricky, as the experiences pictures give rise to are notoriously ambiguous. Imagine looking at a photograph of an apple. If someone asks me ‘What do you see?’, I probably answer ‘An apple’, not ‘A picture of an apple’. This implies that what is experienced when we look at the photograph is ‘an apple’, not ‘a representation of an apple’. At the same time, however, I know very well that I cannot eat the apple on the photograph and that turning the picture around will not reveal the rear side of the apple but only the white rear side of the photograph. This implies that what is present when we look at the picture is not ‘an apple’ – as apples are eatable and three-dimensional – but ‘a representation of an apple’ which is uneatable and two-dimensional. The difficulty for every theory of depiction is that it has to take into account these two contrasting phenomenological facts. When I look at a picture I am, on the one hand, inclined to refer to the *depictum* itself (i.e. the real apple) when someone asks me what I experience. On the other hand, however, I immediately understand that a photograph of an apple is not an apple and I would never confuse them. This structural tension, inherent to every experience of pictures, can be labeled ‘the paradox of pictorial presence’.

In the following, I will not try to solve this paradox. I will rather try to develop an account of depiction that does justice to it. I will defend an anti-representationalist account of depiction according to which pictures allow us to intend the depicted object itself, not its representation. The great challenge such an account faces is that it has to explain the intuitive difference between regular perception and picture-‘perception’. It has, in other words, to account for our direct grasp of the difference between, for example, the real apple and the depicted one. I will develop this account in three sections:

1. **The representationalist account of depiction:** First, I sketch the representationalist account of depiction (Goodman, Scholz). According to this account, pictures do not present me with the depicted thing itself (i.e. apple) but with something different, namely with its representation. Pictures would be a kind of signs which *refer* to the depicted object by means of resemblance. The weakness of this theory is that it can do justice to only one phenomenological fact about pictures: that we know implicitly that there is difference between perceiving an object and seeing it on a picture. It tries to account for this difference by stating that normal perceptions and picture-‘perceptions’ aim for *different intentional objects*. When I perceive ‘an apple ‘in the flesh’, I would experience something different than when I see it in a

picture. This representationalist account is, however, at least one sided as it cannot not do justice to the second basic phenomenological fact about depiction, namely to the everyday practice of referring to the *depicted object* (i.e. the real apple) – and not to the colored piece of photo paper – when asked to describe the intentional object of the experience the picture gives rise to. Whether I actually perceive *apple a*, or look at it on a photograph, I have the experience of intending the exact same *individuum*, and the representationalist theory cannot account for this.

2. **Distinguishing perception and picture-‘perception’:** In this second section, I deal with a question which rises up after my refutation of a representational account of depiction: What about the distinction between regular perception and picture-‘perception’? One could object against my critique of representationalism that if one does away with the idea that perception and picture-‘perception’ have different intentional objects, one cannot account for the distinction between them any more. I will try, however, to account for this distinction *without* appealing to the idea that perceptions and pictures let us intend different objects. I will argue that both allow us to intend the exact same intentional objects, be it in *different ways*: Whereas perception presents its objects as embedded in a causal nexus – the real apple can be eaten, it can rot *et cetera* – pictures can present the same intentional objects but this time as ‘free from causality’. The depicted apple cannot be eaten and it cannot rot, but it is still the same apple as intended in perception. Pictures and perceptions are different, not because they let us intend different intentional object, but because they can present one and the same intentional object in distinct ways.
3. **The picture as a modality of presentation (conclusion):** In the conclusion I argue that one can only account for the ambiguous experience pictures evoke if one adopts a broad concept of experience which, in its turn, incorporates perception, remembrance, imagination, and picture-‘perception’ etc.; all the so-called *modalities of presentation*, all the different ways to be aware of something. When looking at a picture of an apple I do not *perceive* the same thing as when I am looking at a real apple; what I actually *perceive* is something that looks like an apple in certain respects (it is red and has a certain shape) but which also clear differs from it (not eatable, not round, rear side cannot be revealed). I do, however, *experience* the same thing in both cases; I refer to the apple when asked to articulate the intentional object of my experience, not to its representation. Pictures do not, *strictu sensu*, let me perceive the object they depict but they do let me visually experience it, be it in a modus of presentation that differs from the perceptive one: the pictorial modus.

## **Intuitions, Experts and Experimental Philosophy**

Martin Grajner

Institut für Philosophie

Technische Universität Dresden

Traditional philosophical methodology is characterized by its heavy reliance on intuitions. According to the prevalent view, intuitions serve as evidence why certain claims of philosophical interest, like that a particular Gettier-case is not an instance of knowledge, are true. Over the last two decades there have been put forth several challenges to this traditional conception of philosophical methodology. The challenge that I want to discuss in this talk is the challenge of experimental philosophy. My talk is supposed to consist of four sections. In the first section, I want to present the view of intuitions that I favor and that I want to assume in my talk. I advocate a view of intuitions according to which intuitions are understood as seemings, distinct and not reducible to judgments, beliefs or inclinations to beliefs. I want to argue that this kind of view is backed up by certain linguistic data. In the second section of my talk, I want to briefly summarize the main results of the surveys conducted by experimental philosophers. I want to focus mainly on the surveys concerning epistemological thought experiments, that have been conducted by Weinberg, Stich and Nichols (2001). The aforementioned authors argue that intuitions are influenced by such factors as (i) socio-economic status, (ii) cultural descent, (iii) philosophical training and (iv) the order in which the thought-experiments are considered. Weinberg et al. infer from this divergence of intuitions of test persons that intuitions are not reliable and should not be used in philosophical inquiry. In the third section, I want to briefly discuss some replies to the movement of experimental philosophy that I don't believe in. Among these is Ludwig (2007) who has inter alia argued that the persons whose intuitions were surveyed do not possess genuine intuitions in the first place (Ludwig conceives of intuitions as judgments out of conceptual competence). I want to point out that this strand of Ludwig's response does not work, since there is no rationale for denying that the test persons do not possess intuitions. E. Sosa (2007a) has argued that there are different senses of a philosophically interesting concept that are confused by the subjects whose intuitions were tested. I want to argue contra Sosa that this strategy cannot be generalized, since there are cases of conflicting intuitions that cannot be accounted for in this way. In the remainder of this section, I want to present my preferred solution. The solution is, in a nutshell, that primarily expert-intuitions should be trusted in philosophical inquiry. There is prima facie evidence for pursuing this line of response, since most philosophers share the same intuitions when presented with a given thought-experiment. A couple of philosophers have already hinted at or defended such a response to the challenge posed by the studies of experimental philosophers (see for instance Williamson (2007) and Jackson (1998)). But the main challenge for such a view is to characterize the kind of expertise that philosophers possess which makes their intuitions more reliable and which does not beg the question against experimental philosophy. I maintain that there are certain features that philosophers qua experts have that make their intuitions more reliable than those of non-experts. These features are: (i) that philosophers have a higher conceptual competence, (ii) that they are more attentive to certain features of a given thought-experiment, (iii) that their intuitions are not influenced that much by prejudice, misconceptions or emotional attachment, (iv) that they have the ability to back their intuitions by other considerations and (v) that they can explain away deviance in intuitions, especially in cases in which lay persons have diverging intuitions from philosophers. I want to illustrate via case studies how this strategy is to be implemented and how the divergence of intuitions concerning the cases that experimental philosophers have studied, like Kripke's Gödel-Schmidt-case or a Gettier-case,

can be explained away. In the last section of my talk, I want to discuss some objections to this strategy. In a recent paper Weinberg (2009) has argued that a strategy like the one pursued in this talk is not successful. I want to examine some of Weinberg's arguments against attempts to successfully characterize the expertise that philosophers allegedly have. My purpose is to show that they do not harm the view defended in this talk.

#### References:

Bealer, George: „Intuition and the Autonomy of Philosophy“. In: DePaul, Michael and William Ramsey (ed.): *Rethinking Intuition. The Psychology of Intuition and Its Role in Philosophical Inquiry*. London 1998, 200-239.

Earlenbaugh, Joshua and Molyneux, Bernard: „Intuitions are Inclinations to Belief“. In: *Philosophical Studies* 145, 89-109.

Gendler, Tamar: „Philosophical Thought Experiments, Intuitions, and Cognitive Equilibrium“. In: *Midwest Studies in Philosophy* 31/1 (2007), 68–89.

Goldman, Alvin: „Philosophical Intuitions: Their Target, Their Source and Their Epistemic Status“. In: *Grazer Philosophische Studien* 74/1 (2007), 1-26.

Jackson, Frank: *From Metaphysics to Ethics. A Defence of Conceptual Analysis*. Oxford, New York 1998.

Knobe, Joshua und Shaun Nichols (ed.): *Experimental Philosophy*. Oxford, New York 2008.

Kripke, Saul: *Naming and Necessity*. Cambridge (Mass.) 1972.

Laurence, Stephen und Eric Margolis: „Concepts“. In: Zalta, E. (ed.): *Stanford Encyclopedia of Philosophy*. <http://plato.stanford.edu/entries/concepts> (7. 11. 2005).

Ludwig, Kirk: „The Epistemology of Thought Experiments: First vs. Third Person Approaches“. In: *Midwest Studies in Philosophy* 31/1 (2007), 128-159.

Pust, Joel: *Intuitions as Evidence*. New York, London 2000.

Sosa, Ernest: „Experimental Philosophy and Philosophical Intuition“. In: *Philosophical Studies* 132 (2007a). S. 99–107.

Sosa, Ernest: *A Virtue Epistemology: Apt Belief and Reflective Knowledge*. Oxford 2007b.

Weatherson, Brian: „What Good Are Counterexamples?“. In: *Philosophical Studies* 115 (2003). 1–31.

Weinberg, Jonathan, Shaun Nichols und Stephen Stich: „Normativity and Epistemic Intuitions“. In: *Philosophical Topics* 29 (2001), 429-460.

Weinberg, Jonathan (2009): „Are Philosophers Expert Intuitors?“. Typescript.

Williamson, Timothy: *The Philosophy of Philosophy*. Oxford 2007.

## Qué hago cuando creo que $p$

Tobies Grimaltos (Universitat de Valencia) & Sergi Rosell (Sheffield University)

El objetivo general de esta ponencia es profundizar en la comprensión de la naturaleza de la creencia *qua* actitud proposicional, esto es, en tanto que disposición o estado psicológico cuyo contenido es una proposición. En concreto, defenderemos:

- (1) Respecto a la actitud proposicional de creer: que creer que  $p$  consiste básicamente en atribuir una mayor probabilidad subjetiva a  $p$  que a  $\text{no-}p$  (o más exactamente: que a cualquiera de sus alternativas y a todas en conjunto).
- (2) Respecto al contenido: que el contenido de la creencia debe entenderse como la respuesta a una pregunta.

Empezaremos con la defensa de (1) discutiendo las principales alternativas presentes en la literatura.

Una primera posición que consideraremos es la de Carl Ginet (2001), para quien creer consiste en un tipo particular de apuesta. En concreto, creer que  $p$  conllevaría, en primer lugar, *apostar algo a que  $p$  es el caso*. Si bien el mero hecho de apostar algo a que  $p$  es el caso no es suficiente para creerlo, debido a que uno también apuesta algo a que es el caso que  $p$  cuando asume o supone que  $p$ , o cuando se toma a  $p$  como hipótesis. Así, habrá también que *contar con que* es el caso que  $p$ , donde “contar con que” conlleva tanto (a) adoptar una actitud de rechazo respecto de la posibilidad de que  $\text{no-}p$ , como (b) no adoptar ninguna actitud de protección hacia la posibilidad de que  $p$  sea falsa, y así (c) sorprenderse ante la posibilidad de que  $\text{no-}p$ . De este modo, *contar con que* es el caso que  $p$  supone, además de apostar algo por  $p$ , descartar la posibilidad de  $\text{no-}p$ , *sorprendiéndose* si resultara que  $\text{no-}p$ .

Sin embargo, cabe objetar aquí que cuando uno cree que  $p$ , ciertamente se sorprenderá (aunque sea ligeramente) si finalmente es el caso de que  $\text{no-}p$ ; pero sólo si uno está del todo seguro dejará de adoptar protección alguna ante esa posibilidad de que  $\text{no-}p$ , y no siempre (ni siquiera en la mayoría de los casos) se cree con certeza. Creer que  $p$  —por ejemplo, que Nadal ganará a Federer en la final de Wimbledon— no implica creer que  $\text{no-}p$  —que no le ganará— no puede ser el caso en el mundo real; es decir, no implica descartar completamente la posibilidad contraria. Sólo la certeza tiene tal implicación. Además, algunos enfoques equiparan el grado de creencia con lo máximo que el sujeto estaría dispuesto a jugarse en una apuesta que no paga nada si la proposición es falsa y una unidad si es verdadera. Pero la relación entre la confianza, las apuestas y la creencia no es directa, pues se puede apostar por  $p$  en ciertas circunstancias y no en otras, dependiendo de qué esté en juego, con independencia del grado de confianza.

Todo esto nos lleva a rechazar que la apuesta y la sorpresa desempeñen un papel genuino en la definición de la creencia. Creemos, por el contrario, que deberíamos centrarnos en el grado mismo de confianza; concretamente: en la probabilidad subjetiva.

En esta línea, Jonathan Bennett (1990) ha propuesto que creer que  $p$  significa considerar a  $p$  como altamente probable: “creo que  $p$ ” significa “considero  $p$  como muy probable”. La creencia sería, así, cuestión de una alta probabilidad subjetiva. En concreto y como él mantiene, creer que  $p$  sería darle por lo menos un 0.95 de opciones de ser verdadera.

Sin embargo, pensamos que la exigencia de una alta probabilidad es también excesiva. Pues resulta que normalmente creemos cosas a las que asignamos una probabilidad bastante menor. Volviendo al ejemplo anterior, puedo creer que, en la final

del torneo de tenis de Wimbledon que se está disputando entre Nadal y Federer, será Nadal quien finalmente gane. Pero no tengo una seguridad tal como para otorgarle un 0.95 de probabilidad. Por otro lado, el hecho de que otorguemos una alta probabilidad a una proposición, no supone, por sí mismo, que tal proposición sea el contenido de nuestra creencia. Cabe recordar que probabilidad e información pueden estar en una relación de proporcionalidad inversa. Cuanto menos informativa sea una proposición, más probabilidad (a priori) tiene de ser verdadera.

Nuestra alternativa es pues, como adelantamos, la de que creer que  $p$ , en su grado mínimo, consiste en *considerar más probable  $p$  que  $no-p$* ; esto es, considerar más probable  $p$  que cualquiera de las alternativas que supondrían  $no-p$  y que el conjunto entero (la suma lógica) de tales alternativas.

Además,  $no-p$  puede ser un conjunto muy grande de alternativas a  $p$ ; pero uno no necesita saber cuáles son estas alternativas para descartarlas. Si creo que Thomas Mann escribió *La montaña mágica*, estoy descartando, como autores de esta novela, a Hermann Broch, Robert Musil, Franz Kafka, Hermann Hesse y toda la nómina de posibles escritores de *La montaña mágica*, que no necesito saber quiénes la integran. Qué supone  $no-p$  (qué estados de cosas implican  $no-p$ ) dependerá de a qué pregunta intente responder  $p$  —y esta es la segunda tesis (complementaria de la primera) que tratamos de defender.

La misma proposición puede ser una respuesta a diferentes preguntas y, quizá, qué creencia sea (qué contenido tenga) dependa de a qué pregunta pretenda contestar y qué otras posibles respuestas se descarten. Del mismo modo, el grado de creencia quizá depende del grado de exclusión de las otras posibles respuestas (hasta qué punto hayan sido excluidas). Así, por ejemplo, “Thomas Mann escribió *La montaña mágica*” puede ser tanto una respuesta a “¿Qué famosa novela escribió Thomas Mann?” [(Qué  $x$  tal que  $x$  es una famosa novela) (Thomas Mann escribió  $x$ )] como a “¿Quién escribió *La montaña mágica*?” [(Qué  $x$  tal que  $x$  es un escritor) ( $x$  escribió *La montaña mágica*)]. Y uno puede creer que Thomas Mann escribió *La montaña mágica* en distinto grado dependiendo de a qué pregunta pretenda contestar, pues los valores de  $x$  serán distintos, y puede descartar unos u otros también en distinta medida. Esto es así, además, porque dependiendo de a qué pregunta se pretenda responder, las evidencias o las razones de las que dispondrá y que utilizará un sujeto serán distintas y son éstas las que determinarán su grado de certeza.

Las preguntas a que se pretende responder no sólo determinan, pues, la creencia que se tiene, en su grado y, en parte al menos, su contenido, sino que también orientan la indagación y de qué razones se servirá el sujeto. Del conjunto global de las creencias del sujeto sólo algunas son consideradas por él como razones para creer que  $p$ . ¿Qué es lo que hace que una creencia determinada sea tomada como razón para creer? Nuestra posición es que son razones del sujeto para creer que  $p$  aquéllas que el sujeto considera que hacen probable que  $p$ , en el sentido concreto de que *hacen más probable  $p$  que  $no-p$*  (o que cualquiera de las alternativas a  $p$  que impliquen  $no-p$  y que esté al alcance del sujeto considerar). Puesto que creer que  $p$  es admitir la verdad de  $p$ , es constitutivo de considerar que algo es una razón a favor de  $p$  que el sujeto piense que esa razón, usualmente en compañía de otras razones, hace probable la verdad de  $p$ . Así, no podemos tomar una creencia como razón para creer si no pensamos que tal vínculo se da.

## Motivations for Relativist Solutions to Disagreements

Steven D. Hales  
Department of Philosophy  
Bloomsburg University  
Bloomsburg, PA 17815  
USA  
hales@bloomu.edu

Relativism is a way to resolve disagreements. Given that there are several ways to address disagreements, what is sufficient motivation to adopt the relativist approach? In this talk I will lay out the case for what kinds of epistemic clashes offer the strongest motivation to vote for relativism. I will argue that the most promising candidate on the relativist ticket is that of disputes involving irreconcilable differences. Genuine irreconcilable differences are scarce, and I will argue that the usual proposals of relativism-motivators, such as predicates of personal taste, fail to generate them. I will then argue that irreconcilable differences are to be located at the level of independent methods of generating noninferential beliefs which are then used as basic data for building theories that one holds in reflective equilibrium. Epistemic failure at the meta-evidential level prompts relativism.

One of the longstanding objections to relativism is that there is no reason to prefer it to alternate approaches to settling disputes, approaches that do not carry the controversial baggage of possible incompatible truths. Some of the most salient alternatives to relativism include:

*Keep arguing until your opponent capitulates.* Continued argument is the default mode, not just in philosophy, but continued debate until widespread consensus is reached is the common course of action. For philosophers who assume there is an absolute truth that we can ferret out through universally recognizable reason, it is easy to see the appeal of the relentless argument approach.

*Compromise.* Other times it seems that two parties have dug in their heels to an extent that each might despair of ever bringing their opponent around to their point of view. In such cases it may not be clear what kind of evidence would definitively decide the issue for one side or the other. When that happens, compromise starts to look attractive. Compromise is a bit more contentious than fighting until capitulation, and it may not appeal to someone antecedently convinced that they are in possession of the complete truth on a topic

*Uncover an ambiguity.* A dispute may seem intractable because the two parties are talking past each other and using their words with different meanings. Once the disputants work out their linguistic differences and settle on a common usage, they can reach agreement. A second philosophically current form of ambiguity is contextualism. According to contextualism, the meaning of a term fluctuates between or among different senses, depending on features such as conversational context, standards of utterance, and so on. John MacFarlane (2009) notes that many object to contextualism on the grounds that contextualism solves disputes by showing that there never was a true disagreement at all; according to the detractors, this fails to respect the intuition

that any adequate solution to disagreement ought to allow that there really was a dispute to start with. Therefore contextualism is not a satisfactory solution.

*Agree to disagree.* When arguments have gone on a long time without any realistic prospect of resolution, neither party is willing to compromise, and neither finds an appeal to ambiguity or contextualism satisfying, then they may decide to agree to disagree. Agreeing to disagree solves a dispute by giving up the fight, and has a respectable philosophical pedigree in the Pyrrhonian scepticism of Sextus Empiricus. The Pyrrhonian skeptic aims for peace of mind (*ataraxia*) through the indefinite suspension of judgment (*epochē*).

Given that there are alternatives to adopting relativism, the subsequent task is to discover what sorts of disagreements, are best solved by these familiar strategies, and whether there are any disagreements for which the more controversial relativist approach is the most promising. My thesis is that relativist solutions are most plausible when there is meta-evidential disagreement; i.e. disagreement about what would even count as legitimate first-order evidence to settle a dispute.

Given that we have a variety of options for addressing disagreements, when should we vote for relativism as the best solution? The most promising idea is that there are certain kinds of disagreements to which continued argument, compromise, ambiguity, and Pyrrhonism give especially unappealing answers. The best place to look is at disagreements that in some manner seem epistemically irresolvable; otherwise continued argument is presumably the default approach.

A popular place to look for irresolvable disagreements is with predicates of personal taste, a view defended by Crispin Wright and Max Kölbel. I will argue that differences in taste insufficiently motivate relativism, and that there are plausible moves to be made by taste absolutists. When disputing parties agree as to what counts as basic evidence, there is room for non-relativist solutions to their disagreement. Two eyewitnesses might disagree about what they saw, and be unable to convince the other, but nevertheless they agree that the proper sort of evidence is visual. They disagree about the data, but not about the *kind* of data that is relevant. The same situation holds for matters of taste. Two parties who disagree about whether rhubarb is delicious agree that the information relevant to its deliciousness (or lack thereof) involve the taste, smell and texture of rhubarb. It was just that, like the eyewitnesses, they disagree about the data. Yet it is entirely reasonable to hold that there is an absolute truth of the matter and the disagreements arise out of the disputing parties possessing incompatible, but incomplete, evidence.

When a dispute concerns what even counts as basic evidence (and hence appropriate methods of getting evidence) for the subject matter at hand, then relativism begins to look like the only viable candidate for settling the disagreement. I will argue that only the following sort of disagreement will secure relativism.

1. disagreement over some claim P
2. disagreement over what kind of evidence is relevant to settling (1)
3. disagreement over what kind of evidence is relevant to settling (2)

I will argue that it is the third step that is the relativist's trump card.

## **Perspectiva fenomenológica y el inevitable sentimiento de libertad**

María del Rosario Hernández Borges  
Facultad de Filosofía, Universidad de La Laguna, Campus de Guajara, La Laguna,  
38071, Tenerife  
[roherbo@ull.es](mailto:roherbo@ull.es)

Las diversas respuestas que, desde la filosofía, se han dado al problema de la voluntad libre y la responsabilidad moral han tenido siempre que considerar que, más allá de los argumentos, en nuestras decisiones y en nuestras evaluaciones morales los agentes nos sentimos libres. Seamos compatibilistas, libertarios o deterministas, nuestra explicación de la creencia en la voluntad libre ha de dar cuenta de nuestro sentimiento de libertad.

Siguiendo esta idea, mantendré que nuestra perspectiva fenomenológica, nos sitúa en un lugar desde el que los resultados de los procesos mentales (creencias, deseos, intenciones, sentimientos y demás) son considerados como causas no causadas. Este es un proceso natural porque puede explicarse atendiendo a las características de la autoconciencia. Primero, consideraré algunas de estas características para entender por qué tenemos una idea errónea de libertad. Segundo, intentaré explicar cómo, a pesar de que el conocimiento de las causas de nuestras creencias, deseos e intenciones nos hacen tomar una posición determinista acerca de la libertad, no podemos evitar sentirnos agentes libres y formular juicios morales acerca de nosotros mismos y de los demás.

Primero, comentaré dos características de las situaciones de toma de decisión relevantes para este tema:

i) El acceso epistemológico que tenemos a nuestros estados mentales, fenomenológicos o intencionales, nos hacen creer que son privilegiados, entendidos como directos e incorregibles. Conocemos nuestras creencias, deseos e intenciones de forma tan directa como sentimos nuestros dolores. En el momento en que creo, deseo o elijo sé que soy yo quien está creyendo, deseando y eligiendo, no pensamos en nuestras creencias, deseos o elecciones como el resultado de procesos genéticos o educativos, sino como el origen o causa de nuestras acciones.

ii) De acuerdo con A. Damasio (1999, 2003), nuestras decisiones son activadas por las huellas de la experiencia pasada, lo que ha llamado “la hipótesis del marcador somático”. En la toma de decisiones, cuando la situación ha de ser evaluada previamente a la decisión, las reacciones corporales previas en situaciones similares son tomadas como marcas o criterios que dirigen nuestras decisiones. El sentimiento es el resultado de la percepción consciente simultánea de los cambios corporales generados por mecanismos neuronales no conscientes y la situación que provoca la emoción. Ser consciente de que la situación afecta al cuerpo en el momento de la toma de decisión no es lo mismo que ser consciente de porqué esa situación afecta al cuerpo de tal forma y en tal grado. Lo primero se refiere a la metarepresentación de mi cuerpo en este momento y a las elecciones en este momento; lo segundo se refiere a las representación de la génesis del yo en interacción con el mundo externo.

Segundo, la distinción visoespacial entre marco de referencia egocéntrico y marco de referencia aloécéntrico se ha utilizado para distinguir entre representaciones centradas en el agente y representaciones centradas en un estímulo externo e independiente del agente. Esta distinción combinada con la posibilidad de que los

centros de referencia (los objetos representados) sean el propio sujeto agente o los otros se ha utilizado en el ámbito del conocimiento social (Frith & de Vignemont, 2005). Desde esta perspectiva, el objeto de nuestras atribuciones, yo o los otros, pueden ser considerados desde un marco de referencia egocéntrico. Por un lado, se representa al yo de forma egocéntrica cuando la representación deriva de un conocimiento directo del yo, de la propia vivencia o auto-conciencia. Se representa al otro de forma egocéntrica cuando la representación depende de la relación del otro con el yo. Si cambiara esta relación, la representación también cambiaría. Por otro lado, el yo se representa aloécnicamente cuando el conocimiento sobre mí está separado de quien resulta que soy. Se representa al otro de forma aloécnica cuando los estados mentales de los otros son completamente independientes de los propios. Ambos marcos de referencia están conectados mutuamente, aunque no pueden ser reducidos uno al otro: una vida social equilibrada sería el resultado del conocimiento que proviene de la vivencia directa, inmediata y pre-reflexiva de la interacción social regulada por el conocimiento social aloécnico.

Considero que estas distinciones podrían trasladarse al tema de la libertad y la responsabilidad moral para explicar por qué nos vemos libres y responsables moralmente aunque desde una posición determinista esta sea una conclusión falsa. El marco de referencia egocéntrico es el que utilizamos cuando tenemos la vivencia de la propia libertad y cuando vemos a los otros en relación con nosotros mismos, con nuestras propias vivencias. Esa es la perspectiva fenomenológica. El marco de referencia aloécnico es el propio de una perspectiva objetiva, que nos permite representar al sujeto (yo o el otro) a partir de sus aspectos invariantes, como objeto de reflexión independiente del sujeto que lo estudie. Esa es la perspectiva teórica determinista que adoptamos cuando consideramos los estados mentales (los propios y los ajenos) como determinados por factores genéticos o educativos. Sin embargo, y dado que ambos marcos de referencia están conectados mutuamente, nuestro sentimiento de libertad es inevitable.

## Facing some challenges to material conditional

Héctor Hernández O.

IIF-UNAM

[hortiz@filosoficas.unam.mx](mailto:hortiz@filosoficas.unam.mx)

My aim in this work is to propose a pragmatic solution to some challenging problems posed by A. Blum, D. Edgington and others to *Truth-Functional Analysis (TFA)* of the ordinary indicative conditionals. The accusation against *TFA* is that it yields the validity of some arguments which are intuitively invalid. I argue that, at least in the case of these challenges, the actual guilty of this “undesirable consequence” is not the material conditional, but it is rather a wrong formalization of the arguments.

Alex Blum (1986) has objected to treating ‘ $\supset$ ’ as formalization of ‘if ..then...’ of natural language on the ground of an intuitively persuasive counterexample. The example is based on the following theorem:  $(1) (P \& Q) \supset R \vdash (P \supset Q) \vee (Q \supset R)$ .

Blum argues that, although (1) is a theorem of classical logic, it may well be *true* that ‘If both Al and Bill insert their keys, then the vault will open’ and yet be *false* that ‘Either, if Al inserts his key then the vault will open, or, if Bill inserts his key then the vault will open’. However, this example can be understood in two ways: I or II.

(I) A situation where (a) is true, while (b) and (c) are false:

- a). ‘If both Al and Bill insert their keys, then the vault will open’
- b). ‘if only Al inserts his key then the vault will open’ ..... (If  $P$  then  $R$ )
- c). ‘if only<sup>1</sup> Bill inserts his key then the vault will open’ ... (If  $Q$  then  $R$ )

Now, since  $P$  is ‘Only Al inserts his key’ and  $Q$  is ‘Only Bill inserts his key’, then ‘If  $(P \& Q)$  then  $R$ ’ cannot be (a), but ‘if *only* Al inserts his key and *only* Bill inserts his key, then the vault will open’, which is a false conditional because of falsehood of (b) y (c). Therefore, in this case, the counterexample disappears.

(II) Since ‘If  $P$  then  $Q$ ’ and ‘If  $Q$  then  $R$ ’ have to be false in the example considered, it is not sufficient to insert Al’s key nor Bill’s key to open the vault. It is necessary to insert both keys in order to open the vault. In fact, in this context, if we wanted to express the conditional (a) using other words, we would say likely something similar to (d): ‘It is necessary to insert both keys, Al’s key and Bill’s key, to open the vault’. This points out that Blum’s example is actually a conditional of the form ‘*Only if*  $(P \& Q)$  then  $R$ ’ instead of ‘If  $(P \& Q)$  then  $R$ ’.

More reasons in favour of this conclusion are the following: It is clear that neither ‘Not- $B$ ’ follows from ‘If  $A$  then  $B$ ’ and ‘Not- $A$ ’ (a fallacy of *denial of the antecedent*) nor ‘ $A$ ’ follows from ‘If  $A$  then  $B$ ’ and ‘ $B$ ’ (a fallacy of *affirming of the consequent*). However, from ‘Only if  $A$  then  $B$ ’ and ‘Not- $A$ ’ follows ‘Not- $B$ ’ and, similarly, from ‘Only if  $A$  then  $B$ ’ and ‘ $B$ ’ follows ‘ $A$ ’. Now, note that, in the situation II, from (a) in conjunction with ‘Not- $(P \& Q)$ ’ it can be inferred ‘Not- $R$ ’, and from (a) and ‘ $R$ ’ it can be inferred

---

<sup>1</sup> This ‘only’ is necessary, for it is possible that Al inserted first his key and, given that circumstance, it would not be false ‘If Bill inserts his key then the vault will open’, or vice versa when Bill inserted first his key.

‘ $(P \& Q)$ ’. Therefore, the behaviour of (a) is just that of ‘*Only if*  $(P \& Q)$  then  $R$ ’. Thus, Blum’s example is not an instance of (1), but it seems to be an instance of (2): “ $R \supset (P \& Q) \vdash (P \supset Q) \vee (Q \supset R)$ ” which, of course, is invalid not only using ‘if then’ but also using ‘ $\supset$ ’.

Now, let us consider other challenge. Edgignton (1993) has presented the following intuitively invalid argument which, however, when it is symbolized turns to be a valid one:

- |   |   |
|---|---|
| (1) If God does not exist, then it is not the case that |   |
| if I pray my prayers will be answered by Him.           | (1) $\sim E \supset \sim (P \supset A)$ |
| (2) I do not pray.                                      | (2) $\sim P$                            |
| Therefore, God exists.                                  | $\therefore E$ .                        |

First of all, let us consider why the formalization given above is erroneous. Take (1) to mean (1\*):  $\sim E \supset (P \& \sim A)$ , then if God does not exist the speaker *is committed to pray*, but this is implausible, for the speaker, by means of premise (2), discarded that possibility. Intuitively, the speaker does not have to pray in order to turn (1) true in case of God did not exist. For if such were the case, the non-existence of God would lead to a contradiction: “*I pray and I do not pray*”. The truth of “I pray” comes from the truth of the consequent in (1\*): “ $P \& \sim A$ ”, and the truth of “I do not pray” comes directly from the premise (2).

Now, if the non-existence of God really leads to a logical contradiction, then the argument is, in fact, a satisfactory proof of existence of God by *Reduction Ad Absurdum*. Besides, (1\*) is equivalent to “ $(\sim P \vee A) \supset E$ ”, but certainly the speaker does not seek to say that his refusal to pray is sufficient to God’s existence. On the other hand, it is reasonable to admit that the sense of the premise (1) is the following: “If God does not exist, then if I pray my prayers will be unanswered by Him.” The formalization now is “ $\sim E \supset (P \supset U)$ ”, which agrees our intuitions: the speaker is not committed to pray, from these premises follow neither  $E$  nor  $\sim E$ , the premise (1) is equivalent to “ $(P \& \sim U) \supset E$ ” and to “ $(\sim E \& P) \supset U$ ” which are intuitively acceptable.

This simple solution captures our intuitions and, at the same time, it keeps unhurt the *TFA*. Other puzzles can be treated in a similar way because it is often the content (and thus the context) of an indicative conditional of the natural language, rather merely its grammatical form, what determines how it should correctly be symbolized. Of course, the material conditional still has some challenges to overcome, but at least, when the type of arguments considered here are properly formalized to capture the logical structure of the premises and the conclusion, it does not emerge any conflict between ‘If...then’ and ‘ $\supset$ ’.

## References

- Blum, Alex. 1986. ‘ $\supset$  and ‘if...then...’’. *Logique et Analyse* 29, p. 111.  
 Edgignton, D. 1993. ‘Do conditionals have truth conditions?’ *Crítica*, vol.18, pp.3-30.

## How Chance Constrains Credence: Justifying the Principal Principle

Carl Hofer

ICREA Research Professor at UAB

carl.hofer@uab.es

Epistemic agents are often usefully modelled not by (in part) listing their beliefs *simpliciter*, but instead ascribing to them a *rational credence function*, which is mathematically a probability function. Such a credence function takes as argument a set of propositions about which the agent has opinions concerning how likely they are to be true; the output or value of the function,  $Cr(A)$ , from 0 to 1, expresses the agent's confidence in  $A$  in the obvious way.

Such functions are often called *rational credence functions* because we want to insist that they satisfy certain formal constraints, constraints that correspond to various sorts of epistemic goodness that an agent may possess. Foremost among these constraints is the constraint that the credence function should in fact *be* a probability function, that is, satisfy the Kolmogorov axioms (or some other axioms,

e.g. of conditional probability, as long as they satisfy most of our intuitive notions about probability). That rational credence functions should satisfy the probability axioms is usually demonstrated by so-called "Dutch book" arguments.

But satisfying the axioms is a very weak constraint by itself. An agent could, with all the same evidence we have as background, assign credence 0.99 to the proposition that tomorrow the sun will explode; the axioms only insist that he then assign probability 0.01 to the sun not exploding. So philosophers tend to hope that further and stronger constraints on rational credences may be found and defended, constraints that narrow the range of what can be considered "rational" credences for an epistemic agent to have.

One of the most plausible further constraints is what David Lewis dubbed the Principal Principle (PP). PP states a normative connection between *objective probabilities* (or *chances*) and rational credences: roughly, if a rational agent knows that the objective chance of  $A$  is  $x$ , and has no other information relevant to whether or not  $A$  is true, then his degree of belief in  $A$  must also equal  $x$ . Lewis maintained that a theory of objective chance is only acceptable if it is possible, on the basis of the theory, to demonstrate the correctness of PP. Famously, he said he thought he saw "dimly, but well enough" how a reductive Humean account of chance could make such a demonstration possible --but never provided even a sketch of how the demonstration might go. Many philosophers have been skeptical, and maintained either that *no* deduction of PP is possible, or at least that it is impossible for Humean theories (for example, Strevens (1999), or Hall (2004)).

I will discuss two deductions of PP for Humean chance, one first presented in Hofer (2007) and a new argument. If either argument succeeds, it constitutes a powerful consideration in favor of the Humean approach to objective chance.

Before the arguments can be sketched, we need to see the basics of a Humean account of objective chance. Such an account is reductive, and claims that the objective probabilities in our world supervene on the *patterns in actual events* – especially the stable and stochastic-looking patterns discernible in the "outcome" events that can be associated with well-defined "chance set-up" circumstances. One good example would be the patterns of occurrence of the 6 possible outcomes of throwing a die on a gaming table in a casino; another would be the temporal pattern of

decay events in a collection of radioactive atoms of the same kind. The simplest Humean theory of chance is *actual frequentism*: the chance of event type *A* occurring in circumstances *S* is just the number of *A* type outcomes in such circumstances, in world history, divided by the total number of *S* cases, in world history. Actual frequentism is not a good theory of objective chance, and recent Humean accounts “sophisticate” frequentism using Lewis’ Best System approach to lawhood. In brief, the objective chances in our world are the probabilities derivable from a set of (mathematical) probability axioms defined over various types of events or propositions, that together comprise a “best system” of such axioms for our world. What it is to be a *best* system for our world is to have an optimal combination of simplicity (fewness of the basic axioms), strength (amount of events covered) and *fit* (chances suitably close to the actual frequencies, when many events are covered). Crucially, Humean chances need not be exactly the same as actual frequencies, but where the number of events covered by a chance law is high, they must be very close to the frequencies.

The consequentialist argument for PP, in brief, goes like this: If an agent’s credence in the occurrence of *A*-type events in *S*-circumstances is not equal to the objective chance *x*, then at most places and times, the agent’s predictions about the number of *A*’s in (say) 100 *S* trials will either be too high or too low. So there is a clear sense in which the agent “does best” in an uncertain world if his credences match the objective chances.

The *a priori* argument is similar but more formal, and aims to show a logical incoherence in the agent who defies PP. To have credence *r* in *A*-type events in *S* is to commit oneself to the fairness of always betting on *A* in *S* at odds of  $r:(1-r)$ , even in an indefinitely long sequence of *S*-trials. On the other hand, in an *indefinitely* long sequence of *S*-trials, the frequency of *A* outcomes is guaranteed to be extremely close to the Humean objective chance *x*, and hence the only fair odds for betting on *A* repeatedly are  $x:(1-x)$  (or something very close). So if the agent knows the Humean chance is *x*, then if his credence *r* is non-trivially different from *x*, then he believes a contradiction.

Both deductions of PP, I will argue, can be defended against extant objections, and together show that if (and *only* if?) we understand chance along Humean lines, we can demonstrate an important further constraint on rational degrees of belief.

#### *Bibliography:*

Hall, N. (2004), “Two Mistakes about Credence and Chance”. *Australasian Journal of Philosophy*, 82, pp. 93 - 111.

Hoefer, C. (2007), “The Third Way on Objective Probability: A Skeptic’s Guide to Objective Chance”. *MIND* 116, pp. 549-596.

Lewis, D. (1994), Humean Supervenience Debugged. *MIND*, 113, pp. 473-490.

Strevens, M. (1999) “Objective Probability as a Guide to the World”. *Philosophical Studies* 95, pp. 243 - 275.

## **¿Por qué el externismo no es intrínseco a una caracterización teleosemántica del contenido? Un análisis del caso del hombre del pantano.<sup>1</sup>**

Andrés L. Jaume  
Universidad de las Islas Baleares

La defensa del externismo es considerada como un dogma para los proponentes de la teoría teleosemántica del contenido, de suerte que si se es partidario de ésta se es necesariamente externista. Todos ellos (Dretske, Millikan, Neander y Papineau) se basan en el caso estudiado por Davidson del hombre del pantano y asumen que si los contenidos que posee el hombre del pantano son determinados relacionamente entonces aquellos contenidos que resultan fijados a través de la historia selectiva de los mecanismos funcionales que los producen/consumen también tienen una dimensión ineludiblemente externista. Este punto de vista acerca del contenido ha sido contestado recientemente por Braddon-Mitchell y Jackson quienes señalan que la teleosemántica, como teoría que trata de caracterizar el contenido en términos de historia selectiva, es insatisfactoria. Si la historia selectiva es un rasgo esencial del contenido entonces debe haber un contenido determinado para cada historia selectiva. Pero puede suceder que los mecanismos que causan contenidos y que han sido desarrollados evolutivamente sean en realidad los mismos para muchas historias selectivas posibles. Así, argumentan Braddon-Mitchell y Jackson, lo que importa no es la historia selectiva en tanto que historia selectiva, sino que el sistema que maneja las representaciones sea de una determinada clase con la historia evolutiva correcta. Es decir, lo que se busca es la compatibilidad con un determinado entorno. La razón que subyace a este argumento es que la historia selectiva no es una propiedad esencial o constituyente, ya que no afecta a la identidad transmundana. Así, podemos afirmar que un objeto tiene una función sin necesidad de considerar su historia selectiva. Este argumento evidentemente supondría una merma del poder normativizador que atribuyen los partidarios de la teleosemántica a la teoría histórico-etiológica y atendería igualmente contra esta teoría de las funciones biológicas. Los argumentos de Braddon-Mitchell y Jackson han sido respondidos por Papineau y Neander, ambos sostienen que el núcleo central de las teorías teleosemánticas es la afirmación de que los roles de las creencias y deseos se realizan por estados resultantes de procesos de selección en el mundo presente; si hubiera sido otro el escenario selectivo entonces los estados mentales también tendrían una identidad diferente. Es decir, rechazan el argumento modal de Braddon-Mitchell y Jackson y retoman la idea de la constitutividad histórica del contenido.

Mi propuesta consiste en retomar el argumento de Braddon-Mitchell y Jackson, así, sostengo que el argumento del que parten los partidarios de la teleosemántica en su defensa del externismo es defectuoso no solo porque proceda a partir de un concepto erróneo de función biológica que invalida la fijación del contenido, sino porque, como señalan Braddon-Mitchell y Jackson, la historia no es una propiedad esencial de ninguna entidad por lo tanto no es constitutiva ni de la funcionalidad ni del contenido. Para defender esta tesis me baso en que podemos atribuir funcionalidad y contenido sin necesidad de considerar la historia selectiva, únicamente considerando cuál es la situación actual en la que se desenvuelve el organismo. De este modo señalo que el problema al que se ven abocados los partidarios de las teorías histórico-etiológicas es que confunden causación con constitutividad. Dicho esto creo que es posible hacer frente al argumento del hombre del pantano. Por lo tanto, sostengo, la cuestión no es si

---

<sup>1</sup> Este trabajo forma parte del Proyecto de Investigación del Programa Nacional *Puntos de vista. Una investigación filosófica* (Ref: FF2008-01205).

el hombre del pantano tiene contenidos mentales, sino si los llegará a tener. El hombre del pantano y su original Donald Davidson, son dos entidades diferentes, pero nada impide que sus órganos y sus facultades mentales operen de la manera correcta. El hombre del pantano es capaz de aprender como lo fue en su día Donald Davidson, es capaz de todo lo que es capaz Donald Davidson y si su memoria no se guarda fuera de su cerebro es capaz de recordar todo aquello que podría recordar Donald Davidson en las mismas circunstancias. Para ilustrar lo anterior propongo considerar una situación análoga más próxima a modo de experimento mental; preguntémosnos así acerca de qué contenidos poseía la oveja Dolly. A mi juicio la respuesta es la siguiente: exactamente los mismos que su original. Dolly reconoce la comida como comida, tiene miedo cuando el resto de ovejas también tiene miedo, por ejemplo, si se acercara un lobo, nada en Dolly es diferente salvo que Dolly es exactamente una copia de otra oveja que no ha sido obtenida por los medios habituales. ¿En qué se diferencia del caso del hombre del pantano? En nada. El experimento de Davidson ha sido, además, forzado. Davidson lo planteaba únicamente para señalar que las cuestiones relativas a la normatividad del contenido mental no supervenían al individuo, sino que las determinaba el contexto lingüístico y cultural en el que se encontraba.

#### Bibliografía:

- BRADDON-MITCHELL, D. & JACKSON, F. (1997). "The Teleological Theory of Content". *Australasian Journal of Philosophy*, 75(4), 474-489.
- BRADDON- MITCHELL, D. & JACKSON, F. (2002). "A Pyrrhic Victory for Teleonomy". *Australasian Journal of Philosophy*, 80(3), 372-377.
- of New York Press.
- DRETSKE, F. (1981). *Knowledge and the Flow of Information*. CSLI Publications.
- DRETSKE, F. (1986). "Misrepresentation". En Bogdan (ed.) *Belief*, Oxford: Oxford University Press.
- DRETSKE, F. (1988). *Explaining Behaviour. Reasons in a World of Causes*. Cambridge, Massachusets: M.I.T Press.
- DRETSKE, F. (1995). *Naturalizing the Mind*. Cambridge, Massachusets: M.I.T. Press.
- DRETSKE, F. (2000). *Perception, Knowledge and Belief. Selected Essays*. Cambridge: Cambridge University Press.
- MACDONALD, G. y PAPINEAU, D. (Eds.). (2006). *Teleosemantics*. Londres: Oxford University Press.
- MILLIKAN, R.G. (1984). *Language, Thought and Other Biological Categories*. Cambridge, Massachusets: M.I.T. Press
- MILLIKAN, R.G. (1993). *White Queen Psychology and Other Essays for Alice*. Cambridge, Massachusets: M.I.T. Press.
- NEANDER, K. (1995). "Misrepresenting & Malfunctioning". *Philosophical Studies* 79: 109-141.
- NEANDER, K. (1996). "Swampnan Meets Swampcow". *Mind & Language* Vol.11, nº1, pp.118-129.
- NEANDER, K. (2006). *Naturalistic Theories of Reference*, en Devitt, M. y Hanley, R. *The Blackwell Guide to the Philosophy of Language*, Blackwell Publishing, pp.374-391.
- NEANDER, K. (2008). "Teleological Theories of Mental Content: Can Darwin Solve the Problem of Intentionality". En M. Ruse (Ed.), *The Oxford Handbook of Philosophy of Biology*. Oxford: Oxford University Press.
- PAPINEAU, D. (1987). *Reality and Representation*. Oxford: Blackwell.
- PAPINEAU, D. (1993). *Philosophical Naturalism*. Oxford: Blackwell

## Ontological commitment: Syntax, Semantics and Subjectivism

Juan José Lara Peñaranda  
IES. Tirant lo Blanc, Elche  
[juanjoselara@um.es](mailto:juanjoselara@um.es)

It has proved to be an extraordinary difficult question to say not what there is in the world, but merely what there is according to a given theory or discourse. This debate has taken place – and it still does – under the long shadow of Quine’s ontological criterion (QOC). This says that to be is to be the value of a bound variable. I. e., in order to know the ontological commitments of a discourse, this must be regimented in first-order logic and then, the entities the quantifiers range over are the entities ontologically assumed by the discourse.

I would like to stress three points here:

- (a) QOC could only work for a language in which *all* terms are referring terms, and this is highly unusual, in ordinary talk as well as in science. Quine, conscious of this problem (s., specially Quine (1969a), emphasizes the fact that it is not terms that carry ontological commitment, but existentially quantified terms (e.g. “there are electrons”). He says: “I deplore the philosophical double talk, which would repudiate an ontology while simultaneously enjoying its benefits” (1960, p. 242). But this response is not satisfactory. Nothing prevents us from prefixing a quantifier to non-referring terms; and, obviously, if a term is considered to be non-referring, this quantification will also be considered as such. It is a plain empirical fact that there are instrumentalist scientists and nominalist speakers who form quantifications with non-referring terms (think of van Fraassen or Niels Bohr, or most of us when saying things like “there’s elegance here!” when entering a Viennese coffee house or reading a Quine’s article. Or think of nominalists about numbers, as J. Azzouni).
- (b) Another possible move: when doing science we are trying to describe reality and all our statements must therefore be considered literally; all terms must therefore be referring terms. (Remember the “I deplore” quotation above). The problem with this move is that it starts from a presupposition about what it is to be a literal assertion. This move – call it the “paraphrases move” (PM) – needs to have a criterion to offer literal paraphrases of all the sentences in a discourse, but in order to offer these paraphrases we need to know the ontological commitments of the discourse, which was what we were searching for. Another problem with PM is that it infers that all terms are referring terms from the fact that the sentences are to be taken literally, but this is a *non sequitur*. It is perfectly coherent to consider a given assertion the most literal way we have to describe a fact and the assertion containing non-referring terms.
- (c) In fact, Quine’s criterion could not work even in the case where all terms were referring because of inscrutability of reference (IR). As it is well-known, IR is the idea that we could assign different entities as denotata of the terms without varying the truth-conditions, and hence the meaning, of the embedding sentences. Proxy functions are a simple proof of IR; these are “one-to-one reinterpretations of

objective reference. They leave the truth values of the sentences undisturbed.” (Quine, 1995a, p. 72). But then, it does not make sense to search for ontological commitment in the entities the terms (the bound variables) have to be construed as ranging over, as they can be construed as ranging over a number of completely different ones.

What these considerations show is not that Quine’s exposition of the criterion is in need of correction but rather the impossibility of a *syntactical* criterion. If not syntax, what about semantics? The ontology of a discourse, after all, must be built into its semantics; must it not? There is a tempting Quinean line of reasoning here.

Consider semantics as the theory about the truth-conditions of the sentences of a language (call this thesis S:T-C). This theory will surely use notions as reference – and give an ontology along with it – , but only in order to offer the truth-conditions of the sentences of the language. “Reference and ontology recede thus to the status of mere auxiliaries [of the metalanguage]” [Quine (1990a), p. 31].

I call this idea *ontological deflationism* (OD), i. e., the thesis that languages do not have an ontology of any kind – ontology is a device of metalanguages to offer truth-conditions. (Distinguish this idea from IR: *terms* do not carry determinate ontological commitment).

OD must be false, as it seems absurd to claim that speakers or scientists do not commit themselves with certain entities and not with certain others. The question then is: does S:T-C imply OD? My idea now is that it does in a sense and it does not in other.

The fact that the truth-conditions (understood as sentence application pattern or assertibility conditions) of the sentences of an instrumentalist scientist or a nominalist speaker are exactly the same as the ones of the realist scientist or Platonist speaker implies that the ontology cannot be read off of the semantics of a language. In this sense, S:T-C implies OD. But this does not mean that scientists/speakers do not take on ontological commitments in their discourses, it only means that these are not to be found at the level of truth-conditions. But, if ontological commitment is not in syntax nor in semantics, where is it? Ontological commitment must be seen as a pragmatic question. It may even imply an uncomfortable mentalism regarding ontology, but surely we do not want to give an incorrect answer to a question just because it carries mentalism and therefore subjectivism along with it.

#### REFERENCES:

- Quine, W. V. O. (1960), *Word and Object*, Cambridge, Mas., The MIT Press.  
---- (1969a), “Existence and Quantification.” In Margolis, J. (ed.) *Fact and Existence*. Oxford: Blackwell, pp. 1-17. Reprinted in Quine (1969b), pp. 91–113.  
---- (1969b), *Ontological Relativity and Other Essays*. New York: Columbia University Press.  
---- (1995), *From Stimulus to Science*, Cambridge, Mass.: Harvard University Press.

## Davidson, metáfora y dualidad\*

José Luis Liñán

Universidad de Granada

En “What Metaphors Mean”, Davidson (1978) presenta una línea argumental que ha contribuido a definir el campo de los estudios sobre el lenguaje metafórico, tanto en el ámbito de la filosofía del lenguaje, como en la lingüística o incluso en las ciencias cognitivas o la crítica literaria. Examinaré el argumento y defenderé que su pretendida conclusión no se sigue de los fundamentos que Davidson invoca; más bien, resulta incompatible con las intuiciones sobre la comunicación que el propio Davidson ha hecho valer en otros trabajos. Intentaré mostrar que podemos prescindir de ella, a la vez que conservamos los principios davidsonianos que nos interesan, si adoptamos una concepción *dual* del valor semántico de las preferencias, al estilo de Récanati (2007).

### 1. El argumento

En el argumento de Davidson (1978) conviene distinguir dos tesis:

- (M) *Lo que dicen las metáforas* no difiere del significado literal de las oraciones que las realizan: “a metaphor doesn’t say anything beyond its literal meaning (nor does its maker say anything, beyond the literal, in using the metaphor)” (1978, 246).
- (F) *Lo que dice la mayoría de las metáforas es falso*: “If a sentence used metaphorically is true or false in the ordinary sense, then it is clear that it is usually false” (257).

Según Davidson (1978), *F* es una consecuencia inmediata de *M*. A su vez, *M* puede considerarse una consecuencia del principio de *autonomía del significado* (*A*):

- (A) El significado literal (o convencional) de una expresión es independiente de cualquiera que sea el propósito con el que se profiera, y no puede derivarse de él. “Once a feature of language has been given conventional expression, it can be used to serve many extra-linguistic ends.” (1975, 113)

### 2. El principio de caridad

La tesis *F* no se aviene con el principio de caridad (*C*) que el propio autor invoca en otros escritos (desde 1973), con tal de que constatemos la ubicuidad de las metáforas en el lenguaje cotidiano (*U*). Surge, así, una incompatibilidad entre *F*, *C* y *U*.

- (C) Hay que asignarle significado a las expresiones de una lengua de modo tal que se maximice la verdad y consistencia de lo que dicen sus hablantes.
- (U) La presencia de las metáforas en el lenguaje cotidiano es masiva.

Si *M* es válida, entonces las convenciones lingüísticas están en vigor tanto en los usos metafóricos como en los literales. Tan apropiado es, dadas las circunstancias oportunas, decir “Caen chuzos de punta” como decir “Está lloviendo a cántaros” o “Está lloviendo muchísimo”, cuando efectivamente hay una gran precipitación; en los tres casos, decimos una falsedad si, de hecho, no llueve. No hay ninguna razón por la que los usos metafóricos deban contar como una evidencia de menor categoría o de peor

---

\* Trabajo realizado bajo los proyectos FFI2008-06421-C02-0 (Ministerio de Ciencia e Innovación) y HUM4099 (Consejería de Innovación, Ciencia y Empresa, Junta de Andalucía).

calidad que los literales a la hora de elaborar una teoría de la interpretación que, de acuerdo con *C*, maximice la verdad de lo que dicen los hablantes.

### **3. Una posible vía: el abandono de la autonomía del significado**

Una manera de resolver la incompatibilidad consiste en deshacerse de *F* reinterpretando las tesis *M* y *A*. Davidson (1986) parece dirigirse en esta dirección, atenuando o incluso eliminando el papel de la tesis *A* en la teoría de la interpretación. El precio de abandonar *A* es poner como mínimo en cuestión la empresa misma de una teoría composicional del significado basada en una teoría de la verdad, que Davidson había propuesto en sus clásicos trabajos de los sesenta sobre la aprendibilidad de las lenguas naturales y sobre la relación entre la verdad y el significado, y que están a la base de la propia teoría de la interpretación.

### **4. Dualidad y circunstancias de evaluación**

Ahora bien, no es necesario renunciar a *A*. Considérese la siguiente tesis *D*:

(*D*) *Dualidad*: “To get a truth-value, we need a circumstance of evaluation as well as a content to evaluate” (Recanati 2007, 33).

Incluso si restringimos el contenido al significado saturado de las expresiones —excluyendo así procesos de modulación contextual del contenido—, la tesis *D* desactiva la inferencia de *M* a *F*. Aunque las expresiones usadas metafóricamente signifiquen lo mismo que cuando se usan de manera literal, las distintas circunstancias de evaluación que rigen en cada caso pueden marcar una diferencia en el valor de verdad resultante.

Las metáforas interesantes comportan el uso de expresiones lingüísticas bajo circunstancias de evaluación insólitas o al menos novedosas. Sólo ignorando el papel de las circunstancias de evaluación en la determinación del valor de verdad y de las propiedades inferenciales de las aserciones se puede concluir que las preferencias metafóricas conllevan el empleo de oraciones absurdas o contradictorias. Así nos parecerá cuando contemplemos las expresiones involucradas a la luz de las antiguas circunstancias de evaluación —es decir, justamente cuando no las entendamos.

## **Referencias**

- Davidson, Donald. 1973. “Radical Interpretation”. Reimpreso en Davidson (1984).  
———. 1975. “Thought and Talk”. Reimpreso en Davidson (1984)  
———. 1978. “What Metaphors Mean”. Reimpreso en Davidson (1984).  
———. 1984. *Inquiries into Truth and Interpretation*. Oxford: Oxford UP.  
———. 1986. “A Nice Derangement of Epitaphs”. Reimpreso en: *Truth, Language, and History*. Oxford: Oxford UP, 2005.  
Récanati, François. 2007. *Perspectival Thought*. Oxford: Oxford UP.

## What should we require of an account in normative epistemology?

Robert Lockie

Senior Lecturer, Psychology Dept., Thames Valley University, Paragon House, Boston Manor Rd. Brentford, TW8 9GA, UK.

[bob.lockie@tvu.ac.uk](mailto:bob.lockie@tvu.ac.uk)

The distinction between the regulative and the theoretical pertains to the aims (the desiderata) of an account in normative epistemology. This distinction began in ethics – with the terminology used to make it varying slightly by author. In discussing the relationship between the ethical and the epistemic sense of ‘oughts’, Chisholm (1957) distinguished the *absolute* [objective] from the *practical* [subjective], noting this distinction in Richard Price. Plantinga (1993) noted this distinction in Aquinas; Owens (2000) noted it in Sidgwick. In epistemology, the distinction was also developed significantly by Alston (1985). In ethics direct, the main reviews are in Bales (1971), Brink (1986) and Smith (1988).

The paper notes that internalists and deontologists draw this distinction as a stock response to a stock argument levelled against them by externalists and consequentialists. This argument is that the *Ought Implies Can* principle that is internal to deontology/internalism requires that deontic/internalist justification be perspectively attenuated (Chisholm’s and Alston’s ‘subjective’ sense of right – the classic accessibility requirement of internalism). The externalist/consequentialist objects, however, that an agent may be subjectively, perspectively, as justified as they can make themselves – yet still be objectively wrong. The stock response to this argument by the internalist/deontologist is to cede to their opponents the ‘objective’ (absolute, theoretical) desideratum of a theory of justification; yet take their task to be articulating a core notion of justification that is agent-centred, practical, action-guiding – an account that can serve to regulate the agent’s thought and action.

The paper notes a parallel argument that in turn motivates externalists/consequentialists to embrace this distinction in the teeth of a stock counter to their position by internalists/deontologists. This stock counter is that externalist accounts cannot offer us a ‘doxastic decision procedure’ (Goldman 1980). Standardly, externalists just disavow any claim to offer such “rules for the direction of the mind” but in doing so they just are drawing a regulative/theoretical distinction – taking their task to be satisfying the latter desideratum alone.

The two families of theory, internalism and externalism, have then, as an outcome of their conflict, each been forced to draw the same distinction: a distinction between two separate desiderata of adequacy for an epistemic theory. Under severe pressure from the other’s arguments, each has been forced to abandon any pretensions to satisfy one arm of that distinction - to meet one desideratum of adequacy for an account in epistemology. Between them, they account for both desiderata, separately, they each account for one alone. It is suggested that this situation could be seen as permitting us to finesse much partisan conflict between accounts of justification in normative epistemology (and ethics). It is argued that each of internalism, externalism is essential

to a complete account of justification in epistemology, but that these species of justification answer to radically different desiderata of adequacy for any such account. Finally, and most contentiously, it is suggested that we should see knowledge as needing to answer to the theoretical desideratum of adequacy alone; and rationality as answering to the regulative desideratum of adequacy alone. Such a division of labour may make space for Gordian [knot] accounts. It will, for instance, licence our choosing to develop theories of knowledge that wilfully disown any requirement that a theory of knowledge attempts to answer to the regulative desideratum of adequacy. This will make space for accounts that are at least in the vicinity of Sartwell's (1991, 1992) neo Theatetan position ('knowledge as merely true belief'). The arguments here are assimilated to cognate positions and arguments currently and historically under discussion in the literature (Zagzebski's (2003) 'espresso machine' analogy, the reawakening of interest in Plato's *Meno* problem concerning the value of knowledge over true belief, and the 'Swamping Problem'). Such accounts will be severely jarring to our ordinary language intuitions; but it is argued that in developing an epistemological axiology, data from any such intuitions is not to the point.

## References

- Alston, W. (1985) 'Concepts of epistemic justification', *Monist*, 68.
- Bales, R.E. (1971) 'Act-Utilitarianism: account of right-making characteristics or decision-making procedure?' *American Philosophical Quarterly*, 8, (3).
- Brink, D.O. (1986) 'Utilitarian morality and the personal point of view', *Journal of Philosophy*, 83, (8).
- Chisholm, R. (1957) *Perceiving: A Philosophical Study*, Ithaca: Cornell.
- Goldman, A. (1980) 'The internalist conception of justification', *Midwest Studies in Philosophy*, Vol. 5: *Studies in Epistemology*, eds French, Uehling & Wettstein, University of Minnesota Press.
- Owens, D. (2000) *Reason Without Freedom*, London: Routledge.
- Plantinga, A. (1993) *Warrant: the Current Debate*, Oxford: Oxford University Press.
- Sartwell, C. (1991) 'Knowledge is merely true belief', *American Philosophical Quarterly*, 28, (2).
- Sartwell, C. (1992) 'Why knowledge is merely true belief', *Journal of Philosophy*, 89, (4).
- Sidgwick, H. (1907) *The Methods of Ethics*, London: Macmillan.
- Smith, H. M. (1988) 'Making moral decisions', *Nous*, 22: 89-108.
- Zagzebski, L. (2003) 'The search for the source of epistemic good', *Metaphilosophy*, 34, (1-2): 12-28.

## Indifferentism in Meta-Metaphysics

Dan López de Sa

ICREA Researcher at Universitat de Barcelona

LOGOS-Grup de Recerca en Lògica, Llenguatge i Cognició

Montalegre 6, 08001 Barcelona

dlopezdesa@ub.edu

In the last couple of years, there has been a renewed interest in issues in meta-metaphysics on the nature of certain apparent disputes in metaphysics.

The underlying worry seems to be that some of them are *merely apparent disputes* in metaphysics. The worry is sometimes voiced by means of locutions such as 'merely verbal issues,' 'no (ontological) fact of the matter,' disputants being 'just talking past each other' and 'not substantially disagreeing,' using 'different ways of talking about the same,' and, it seems fair to say, the worry comes in general with a "deflationary," "dismissive" attitude with respect to the apparent dispute in question.

In my view, such a talk typically conflates two quite radically different ways in which it can be held that an apparent dispute in metaphysics is merely so. On the one hand, it might be that the dispute is indeed genuine, but of a semantic rather than metaphysical character. On the other, it might be that the apparent dispute is just merely apparent, given that the views allegedly under dispute turn out to be, in a certain sense, equivalent to each other.

The main aim of this paper is to taxonomize these different attitudes, and in particular to characterize the latter by contrast with the former, which I have labeled *indifferentism*.

I will offer a toy example where indifferentism seems clearly vindicated, in terms of the "dispute" between the Top theorist, according to which at the fundamental level of reality there are facts about things being on top of other things and facts about things being underneath other things are merely derivative, and the Bottom theorist, according to which the situation is the other way around. I will use this case to assess indifferentism with respect to philosophically interesting apparent disputes in metaphysics, notably the debate on the persistence of ordinary objects over time between endurantism and perdurantism.

Indifferentism proper will be contrasted with cases where the appropriate attitude seems to be to acknowledge that certain apparent disputes in metaphysics are indeed genuine disputes, but it turns out that they are of a semantic nature. I will illustrate this with the debate between so-called "supervaluationist" and "almost-identity" solutions to the problem of the many.

(Both attitudes are in contrast with that which regards a dispute as both genuine and metaphysical in character. Sometimes this is explicated in terms of disputes which are not answerable to considerations of "analyticity" or "conceptual analysis." A subsidiary aim of this paper is to suggest that this need not be so.)



## Idealistic Moments in Late Wittgenstein

Jakub Mácha, Dept. of Philosophy, Masaryk University, Arna Nováka 1, 602 00 Brno, Czech Republic, macha@mail.muni.cz

“Hegel seems to me to be always wanting to say that things which look different are really the same. Whereas my interest is in showing that things which look the same are really different.”

Wittgenstein

It is tempting to think that at the outset of analytic philosophy Russell and Moore revolted against the neo-Hegelian movement by arguing in favor of pluralism, realism, common sense, and by defending the reality of external relations. A relation is *internal* to some of its terms for if the relation were taken away from it, then the term would no longer be the same thing. We can further define that a relation is *internal (per se)* if it is internal to all its terms, otherwise it is *external*. We could therefore say that an internal relation is *essential to its relata*.

An extreme position has been proposed by the neo-Hegelians (e.g., by Bradley, Royce or Joachim). This is that all relations are internal. This position is called *the doctrine of internal relations* and results in idealism and monism. Although Russell and Moore rejected this line of argument, some recent investigations indicate that it may be sound (see Schaffer, forthcoming).

The terms or *relata* in question were mostly particulars; the problem becomes more complicated if one takes universals into account (cf. Rorty 2006: 335–6) as Wittgenstein did. Wittgenstein himself was not immediately involved in this debate, even though there is an echo of it in his *Tractatus*. The notion of an internal relation is, however, central to this book. According to Wittgenstein there is the “internal relation of depicting that holds between language and the world” (4.014). Furthermore, he declares he has “the answer to the vexed question ‘whether all relations are internal or external’” (4.1251). There are, however, two or three approaches to interpreting the *Tractatus* and the differences among them can be shown, *inter alia*, with respect to the notion of internal relation (cf. McGinn 2009:1). In the first interpretation, which is called “metaphysical” and advocated notably by P. Hacker (2001), the internal relation is a substantial relation of isomorphism between language and the world. Only such a sign that is standing in this internal relation to a possible object expresses a sense. The relation is internal to the sign, for if it didn’t have it, it wouldn’t be any sign at all. In the second main approach (entitled “anti-metaphysical”; see, e.g., McGinn 2009), the internal relation of depicting holds within the language. A way of symbolizing stands in this internal relation to a symbol or, in other words, a way of symbolizing is grounded in the nature of a symbol (cf. 5.473: “Logic must look after itself. If a sign is possible, then it is also capable of signifying.”) The third (so called “resolute”) interpretation says that the concept of the internal relation of depicting is inherently incoherent and has only a transitional role. It is the ladder that must be thrown away (see Diamond 2002: 70). An internal relation is strictly speaking no relation at all, and hence Wittgenstein regards only external relations as proper (4.122).

We can see that Wittgenstein had significantly employed the concept of an internal relation, and at least in the first and second interpretation he admitted the reality of internal relations without advocating the doctrine of internal relations. In his late phi-

osophy of psychology there is, however, a remark which may give the impression that Wittgenstein came back to the doctrine of internal relations. This remark reads:

One might even feel like this: “Everything is part and parcel of everything else” (internal and external relations). Displace a piece and it is no longer what it was. Only in this surrounding is this table this table. Everything is part of everything. [I believe Hegel meant something like this.] Here we have the inseparable atmosphere. (*Remarks on the Philosophy of Psychology* I, § 339)

This remark belongs to a larger sequence about the concept of an *atmosphere*, in which Wittgenstein intended to capture the mental or psychological dimension of our language. An atmosphere of a word is “a ‘corona’ of lightly indicated uses” (PI II, iv). What also could Hegel mean here (this parenthesis occurs only in an earlier manuscript)? In the very first section of the *Phenomenology of the Spirit*, Hegel argues that “a this Here [...] is *not* this Here, but a Before and Behind, an Above and Below, a Right and Left” (p. 64). The “Here and Now” or the “Thinghood” is constituted by relations to its surrounding and thus “it is only the sense-certainty as a *whole* which stands firm within itself” (p. 62).

I will argue, that according to the metaphysical interpretation above, the late Wittgenstein can be seen as an adherent of the doctrine of internal relations and as a monist (more precisely as a *world-first* monist in the terminology developed in Schaffer 2010). On this reading, every single thing is internally related to every other thing in the universe.

In the anti-metaphysical approach, not the universe, but language is to be regarded as a unity. The “corona”, so to speak, can cover the whole of language. What is then the internal relation in question? The late Wittgenstein mentioned an internal relation in the connection with the phenomenon of seeing-as of an aspect, in order to blur the difference between thinking and seeing: “what I perceive in the dawning of an aspect is [...] an internal relation between it and other objects” (PI II, xi). When a thing A is *seen as* some other thing B, then there is an internal relation between A and B perceived and thought of. All two things would be, thus, internally related as long as one is seen as and thought of another, i.e., they can be named phenomena of the same subject or—in Hegel’s terms—of the same consciousness.

I shall conclude by maintaining that the metaphysical approach brings Wittgenstein nearer to Bradley et al., whereas the anti-metaphysical nearer to Hegel.

## References

- Diamond, C. (2002), ‘Truth before Tarski’, in E. H. Reck (ed.), *From Frege to Wittgenstein: Perspectives on Early Analytic Philosophy*. Oxford: Oxford University Press.
- Hacker, P. M. S. (2001), ‘Naming, Thinking and Meaning in the *Tractatus*’, in *Wittgenstein: Connection and Controversies*. Oxford: Oxford University Press.
- Hegel, G. W. F. (1998), *Phenomenology of Spirit*. Translated by A. V. Miller. Delhi. McGinn, M. (2009), ‘Wittgenstein and Internal Relations’, *European Journal of Philosophy* (early view), pp. 1-15.
- Rorty, R. (2006), ‘Relations, Internal and External’, in *Encyclopedia of philosophy*, 2<sup>nd</sup> Edition, D. M. Borchert (ed.), Detroit: Thomson Gale, vol. 8, pp. 335-345.
- Schaffer, J. (2010), ‘The Internal Relatedness of All Things’, *Mind*, forthcoming.
- Wittgenstein, L. (1995), *The Collected Works of Ludwig Wittgenstein*. Oxford: Blackwell.

## It is not presuppositions which are cancellable

Josep Macià

LOGOS / Universitat de Barcelona. josep.macia@ub.edu

Providing a proper characterization of what presuppositions are should help elucidate the nature of meaning and of the different levels, aspects or kinds of meaning (asserted meaning, literal meaning, character, content, conversational implicatures, explicatures, presuppositions, conventional implicatures, expressive meaning, etc.). There is a promising view of presuppositions which treats them as semantic entailments, and which models them using partial meaning functions. This promising view is challenged by the view put forward by those authors (for instance Soames(1989), Green(2000) and Potts(2004)) that argue that presuppositions (like conversational implicatures) are cancellable. I will argue that the data that has been used to argue for this view can be more appropriately accounted for assuming the semantic view of presuppositions. I will argue, therefore, that we have no good reason to abandon the simpler, clear, promising semantic view according to which presuppositions are not (in the relevant sense) cancellable.

- Conversational implicatures are cancellable, as illustrated in (1); logical and analytical implications are not cancellable, as illustrated in (2)

(1) I like some of your friends. And I do not mean to say that there are some friends of yours that I don't like; I like all the ones I know, and I am sure I will also like the others when I meet them

(2)\* Ann is a sister of Beth. And I do not mean to say that Beth is also a sister of Ann

- Presupposition is the kind of entailment relation illustrated in (3)-(5)

(3) John has stopped smoking → John used to smoke

(4) It was in Boston where Joe met Quine → There is a place where Joe met Quine

(5) The king of France is wise → France has a king

Unlike logical implications and asserted content, presuppositions are propositions that are "taken for granted".

- Partial function treatment of presuppositions: The meaning of an expression is a function that for each context assigns a value. If an expression carries a presupposition the function is a partial one. For instance:

$[[stop]](w,t)(a,P) = \begin{cases} \text{True} & \text{if } Pa \text{ in } w \text{ at some time } t' \text{ earlier than } t, \text{ and not-}Pa \text{ in } w \text{ at } t \\ \text{False} & \text{if } Pa \text{ in } w \text{ at some time } t' \text{ earlier than } t, \text{ and not-}Pa \text{ in } w \text{ at } t \\ \text{Not defined} & \text{otherwise} \end{cases}$

If  $s$  carries a presupposition that  $p$ , then if  $p$  is not the case in the world where  $s$  is interpreted, then  $s$  does not have a semantic value with respect to  $w$ . This semantic view of presuppositions predicts that presuppositions are not cancelable in the way that implicatures are.

- Arguments for the cancellability of presuppositions (Soames(1989), Potts(2005)):

-Cancellation by negation of the implication: The whole sentence (6) does not presuppose that Ali has a brother:

(6) Ali's brother isn't wise, because Ali does not have a brother

Non-inheritance: The whole sentence (7) does not presuppose that Joe met Quine; the whole sentence (8) does not presuppose that France has a king (only that John believes it).

- (7) If Mario met Quine, then it was in Boston where he met him
- (8) John believes that the king of France is wise

- We should distinguish two senses of “cancellable”:

- **Cancellable-1** (“x\_y” denotes the concatenation of x and y)

If an utterance of s in C gives rise to an implication that p, then this implication is cancellable iff there is a text r such that (a) r as uttered in C does not contradict s (as uttered in C), and (b) s\_r as uttered in C does not give rise to the implication that p

Notice that this sense of “cancellable” would not allow us to say that there is “cancellation” of presuppositions in some “non-inheritance” cases such as (8) (since, (8) does not consist in just adding some text to the end of the first sentence in (5)).

- **Cancelable-2**

If an utterance of s in C gives rise to an implication that p, then this implication is cancellable iff there is a text T such that: T contains s, and T (as uttered in C) does not give rise to the implication that p.

Cancelable-2 is an uninteresting notion of “cancellability”. Even logical implications are cancellable in this sense. For instance, “John and Mary will come” implies “John will come”, but “If John will come, then both John and Mary will come” does not.

The fact that presuppositions are Cancelable-2 does not pose any problem for the semantic view of presuppositions. What needs explaining is rather why in some texts presuppositions are inherited when logical implications are not. We can show that the partial function view of presuppositions allows us to appropriately account for the inheritance and non-inheritance data.

- Are presuppositions cancelable-1? The evidence for this claim is based on examples such as (6) above. My contention is that negation in (6) is metalinguistic (Horn 1989). An utterance of (6) does not express that certain proposition does not obtain, but rather it shows opposition to accepting an utterance of “Ali’s brother is wise”. So in (6) there is no *use* of a sentence that would normally carry a presupposition, and so no cancellation.

- Possible reply to my contention: in sentences (10)-(12) there is cancellation-1, but no negation, and so no metalinguistic negation.

- (10) It is very unlikely that Ali’s brother is wise, because Ali does not have a brother
- (11) I do not think that Ali’s brother is wise, because Ali does not have a brother
- (12) There are very few chances that Ali’s brother is wise, because Ali does not have a brother

- I will argue: (10)-(12) still involve metalinguistic discourse. We can show this by appealing to what I will call the *Fumfling Test*. Notice that (6) and (10)-(12) have analogues which one can easily argue are metalinguistic. For instance:

- (6)b Joe is not fumfling, because *to fumfle* does not mean anything
- (10)b It is very unlikely that Joe is fumfling, because *to fumfle* does not mean anything

## Sellars on the Credentials of Perceptual Knowledge

Edgar Maragat  
Avenida Blasco Ibáñez, 30  
46010 Valencia

1. In his increasingly celebrated critique of epistemic empiricism, Wilfrid Sellars (1956) examines the epistemic authority of observational reports as paradigmatic expressions of empirical knowledge. He aims to give a satisfactory answer to the question: which are the epistemic credentials of reports such as ‘This is a green tie’ when they are uttered in typical situations?

Sellars’ response could perhaps be better understood if we compare it with the most common contemporary reactions to this kind of concern, i.e.:

- (1) Those reports are not in need of credentials (the Wittgensteinian move).
- (2) There are inner episodes that by themselves possess that authority and confer it to those reports (the empiricist move).
- (3) The only credential available is the objective reliability of the reporter as reporter of a particular kind of facts (the externalist reliabilist move).

In his essay Sellars pays no attention to the first move (which would otherwise have cancelled Sellars’ worries), discusses and rejects the second one (in dismantling what he calls the Myth of the Given) and finally finds unsatisfactory the mere appeal to an objective reliability to account for social attributions of knowledge. I will not address here either the rationale for the Wittgensteinian move or Sellars’ argument against the empiricist solution. I will instead focus on Sellars’ reluctance to accept the dominant naturalistic reaction, namely: the attempt to explain knowledge in terms of (reliable) mechanisms of belief formation.

Sellars’ own approach analyzes the source of authority in two factors. Firstly, we have the factual disposition to emit the report or indeed the fact of the report (the thinking out loud ‘This is a green tie’). And, secondly, the knowledge the reporter might have of her own reliability as a reporter in a certain domain. It is the conjunction of a non-inferentially elicited report and the acquaintance of the reporter with her authority as reporter in actual circumstances that credits the report as expressive of knowledge.

This analysis presupposes that the candidate to knower is reliable in forming perceptual beliefs in determinate circumstances, but moreover it emphasizes that she must *know* about her own reliability if we are going to attribute knowledge of the reported facts to her on a justified basis. According to Sellars, if she ignores her reliability, her report is deprived of authority, and when we talk about knowledge, we talk about declarative utterances that enjoy that authority. As other epistemological internalists, Sellars approaches the problem of knowledge as an issue about responsibility in making claims and taking claims for granted.

So, the specific question I will address is this: is he explaining knowledge by appeal to knowledge? Or is he otherwise tempted to explain knowledge without presupposing it?

2. It could seem that Sellars’ point in explaining perceptual knowledge is that only knowledge about many other things may entitle someone to make an observational report that expresses perceptual knowledge (cf. Sellars 1956, §§ 19, 37). This impression appears to be confirmed when Sellars defends his analysis against a plausible charge for circularity just by arguing that to know that we are reliable as reporters in a certain sense, that is, of a determinate kind of facts, we do not need to

assume that in the past we knew that those facts – which some disposition to speak would be a reliable symptom of – obtained. So it seems that Sellars' only worry on this matter concerns the possibility of a *specific* circularity in the argument (i.e. that to know 'This is a green tie' I should have known before that other things were green and those or others were ties).

Nevertheless recent discussions about Sellars' solution (Brandom 2002, McDowell 2009) tend to nuance the import of the second factor of Sellars analysis in accounting for perceptual knowledge. McDowell, for instance, claims that if we construe the second factor – the appeal to knowledge about circumstances, reliability and the like – as a premise of sorts of the perceptual report, we neglect the non-inferential character of the report and risk the ridiculous view that perceptual reports are conclusions drawn from previous knowledge. In McDowell's eyes, even for Sellars the justification of these reports proceeds from suitable experiences (experiences with conceptual content). The knowledge of one's own reliability etc. would only provide a general authority to utter certain reports in the present context. (It could be argued that McDowell aims at giving to Wittgenstein, empiricism and Sellars on empirical knowledge their due at the same time.) Brandom has also some trouble in accepting literally what Sellars says about the second ply of his account. But he complains just about Sellars insisting on the need that the purported knower knows of her reliability. For Brandom it suffices that others take notice of her reliability or that she is recognizably reliable (Brandom 1997, 157 ff. and 1994, 218 ff.).

What those corrections have in common (despite of general disagreement on Sellars' commitment with empiricist and externalist insights) is the warning against exaggerating the epistemic importance of the second factor in Sellars' two-dimensional account. Actually if we were to emphasize the second dimension over the first one as doing the properly epistemic work, we were perhaps precluding the possibility to answer satisfactorily the question that originated the whole debate, since if you are worried about the purported absence of epistemic credentials in behalf of reports such as 'This is a green tie', I suppose that you should feel even more concerned about the credentials of beliefs of the form 'X is a reliable symptom of Y'.

3. I think that if you accept the point of the original question, but you do not feel tempted to a neo-empiricist move (as McDowell's) and you do not want to surrender to non-internalist accounts (as Brandom seems to do and Sellars manifestly tried to avoid), then you should also rephrase Sellars' solution so as to mitigate the impression that some other knowledge supports non-inferential knowledge.

My proposal is what I take it to be the Davidsonian orthodoxy: there are inferential relations between beliefs originated by our causal interaction with the world and other beliefs that we have and although none of those beliefs amounts per se to knowledge, these relations lift some of the caused beliefs (or empirical reports in Sellars' scheme) to the status of perceptual knowledge (or to the status of expressions of knowledge) (cf. Davidson 1983). Of course, this does not provide an explanation of knowledge in non-semantic terms (in full-fledge naturalistic terms), but in accepting this account we get rid of an explanation of knowledge of this basic (or fundamental) kind in terms of knowledge.

## Empirical Data and the Theory of Reference

Genoveva Martí

ICREA & Universitat de Barcelona

genoveva.marti@gmail.doc

E. Machery, R. Mallon, S. Nichols and S. Stich, have argued (*Cognition*, 2004) that there is empirical evidence against Kripke's anti-descriptivist causal/historical picture as a general theory of reference for proper names. Their argument is based on an experiment that compares the intuitions about proper name use of a group of English speakers in Hong Kong with those of a group of non-Chinese American students. In the experiment in question Machery et al. give two groups of subjects (in China and in the US) a story based on Kripke's famous Gödel/Schmidt case:

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

And then they ask the participants:

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

- (A) the person who really discovered the incompleteness of arithmetic?
- or
- (B) the person who got hold of the manuscript and claimed credit for the work?

A significant number of Chinese participants answer (A), whereas Westerners tend to answer (B). This leads Machery et al. to conclude that East Asians are less likely than Westerners to give answers that are in line with the causal/historical picture, something that suggests that the semantic intuitions Kripke appeals to are culture dependent and not universal.

I have contended (*Analysis*, 2009) that the results of Machery et al.'s study do not lead to any relevant conclusion for the semantic theory of names, arguing that their data elicit theoretical judgments, judgments by laymen of how speakers use names. In a recent paper (*Analysis* 2009), Machery and colleagues have presented further evidence to rebut my claim, arguing on the basis of a new study that distinguishes meta-linguistic questions (questions that elicit theoretical judgments) and linguistic questions (which supposedly provide data about use). I wish to argue that the new data presented by

Machery and colleagues does not provide, again, the right evidence for the theory of reference. Machery and colleagues continue to set up their studies by asking questions to participants about how a hypothetical speaker uses a proper name. I will argue that the data collected is still theoretical.

Nevertheless, my criticism of Machery and colleagues opens up a discussion as regards the kind of data that should constitute the input of semantic theorizing. I will argue that, even though I do not think that we should conceive of semantics as a descriptive enterprise, i.e., as a pure articulation of the uses and practices of linguistic communities, the relevant data for semantic theorizing are data about actual use, and the way in which experimental philosophers propose to collect that data is incorrect. The semanticist uses (already does) methods that are closer to the anthropologist's observant participation, and no new methodology of data collection is required to construct semantic theory.

## Negation in self-referential four-valued languages

José Martínez Fernández  
 Department of Logic, History and Philosophy of Science  
 University of Barcelona  
 C/Montalegre 6  
 08001 Barcelona

Let us consider a propositional language, interpreted truth-functionally on a finite set of truth values. A *stipulation* is a system of equations of the form  $p_i = \varphi_i$ ,  $i \in \omega$ , where  $p_i$  are the atomic propositions and  $\varphi_i$  are formulas of the propositional language. Stipulations are meant to represent self-referential nets of sentences. For example, the Liar Paradox ('this sentence is false') can be represented as the proposition  $p$  with the stipulation  $p = \neg p$ ; the Curry Paradox ('if this sentence is true, then the Earth is flat') can be represented as the proposition  $q$  with the stipulation  $q = q \rightarrow r$ ;  $r = 0$ , where 0 represents the truth value false. A stipulation is *consistent* when there is a solution of the system of equations, i.e., an assignation  $v$  of truth values to the atomic propositions such that  $v(p_i) = v(\varphi_i)$ , for all  $i \in \omega$ . Given an interpreted propositional language, we can say that the language has the (*Gupta-Belnap*) *fixed-point property* when every stipulation is consistent. A language with the fixed-point property can evaluate consistently all types of self-referential nets of sentences expressible in the language.

Visser, generalizing previous work done by Kripke, Martin and Woodruff and others, proved (see [3]) that if the operators of a propositional language are monotonic for a ccpo (coherent complete partial order) defined on the set of truth values then the language has the fixed-point property. In [2], Gupta and Belnap proved that the property of monotonicity for ccpo's, although sufficient, is not necessary for the fixed-point property and put the problem of characterizing the fixed-point property. The aim of this talk will be to classify some important four-valued propositional languages with respect to the fixed-point property. From now on, we will assume that all propositional languages we consider include the constant operators (they are convenient to represent in stipulations non-circular sentences as 'the Earth is flat').

It is likely that the most influential four-valued logic is Belnap's logic, which is a natural generalization of strong Kleene logic.

The operators of Belnap's logic are monotonic for the ccpo on  $\{0, 1, 2, 3\}$  that has 2 as least element, 3 as greatest, and 0 and 1 as non-comparable elements (usually called 'the order of knowledge', since it reflects the degree of information provided by the truth values in the interpretation in which the values are epistemic, i.e., when 0 means that there is information in favour of the truth of the sentence, 1 that there is information against the truth of the sentence, 2 that there no information about the sentence and 3 that there is information both in favour of and against the truth of the sentence). There are several interpretations of this logic. We will consider, to fix ideas, that 0 represents the value 'false', 1 the value 'true', and 2 and 3 represent two different ways in which sentences can be pathological. Let us consider the following non-monotonic operators

	$\neg_e$		$\neg_*$		$\downarrow_1$		$\downarrow_1^*$		$\downarrow_2$		$\downarrow_2^*$		$\downarrow_3$
0	1	0	1	0	0	0	0	0	0	0	0	0	0
1	0	1	0	1	1	1	1	1	0	1	0	1	0
2	1	2	1	2	0	2	0	2	1	2	1	2	0
3	1	3	3	3	0	3	3	3	0	3	3	3	1

All of them have natural interpretations:  $\neg_e$  is called exclusion (or strong) negation in the literature. It is obvious that no language can contain that operator without producing a contradiction ( $p = \neg_e p$ ). We propose as an approximation the operator  $\neg_*$ , that we call pseudo-exclusion negation. The operators  $\downarrow_2$  and  $\downarrow_3$  are operators of pathologicity, that reflect in the language the distinctions made in the semantics. So  $\downarrow_2 p$  says that  $p$  has truth value 2. Analogously,  $\downarrow_1$  is an operator of strong truth (or determinate truth). The variants of these operators are designated with a star.

All these operators are non-monotonic for the order of knowledge, so we can ask how many of them can be added to the Belnap logic while preserving the fixed-point property. The solution is given by the following results (which are easily proven from well-known results in clone theory):

**1 Theorem.** Belnap's logic (with constants) is maximal for the fixed-point property.

**2 Corollary.** None of the previous operators can be added consistently to Belnap's logic.

Since Belnap's logic cannot be expanded, our following aim is to consider a four-valued generalization not of the strong Kleene logic, but of the weak Kleene one. Among the possible generalizations we will focus on this one, studied by Fitting in [1]:

	$\neg b$	$\wedge_{sw}$	0	1	2	3
0	1	0	0	0	0	3
1	0	1	0	1	2	3
2	2	2	0	2	2	3
3	3	3	3	3	3	3

First let us present a negative result:

**3 Proposition.** The language generated by the constants and the operators  $\neg b, \wedge_{sw}, \downarrow_2$  does not have the fixed-point property.

Proof:  $x = \neg b \downarrow_2 (2 \wedge_{sw} x)$ .

Then we will present the following positive results:

Let us consider the set  $E_k = \{0, 1, \dots, k-1\}$ . Given an interpreted language  $F$  on the set of truth values  $E_k$ , let us call  $F^{(1)}$  the set of all the unary operators expressible in the language and  $F^{(1-1)}$  the set of functions in  $F^{(1)}$  which are permutations. For  $a \in E_k$ , the stabilizer of  $a$  (denoted as  $st(a)$ ) is the set of all permutations  $f : E_k \rightarrow E_k$  such that  $f(a) = a$ .

Given a partial order  $(E, \leq)$ , let us call  $Mon(\leq)$  the set of all functions monotonic on  $\leq$ . The flat ccpo on  $E_{k+1}$  is the partial order  $\leq_k$  defined by  $k \leq_k i$  for all  $i \in E_{k+1}$ .

Let  $f$  be a finitary function on  $E_{k+1}$ . The derived set of  $f$ , denoted  $der f$ , is the set of all functions which can be obtained from  $f$  with some (all, none) of its variables replaced by constants. The restriction of  $f$ , denoted  $re f$ , is the function  $re f : E_k \rightarrow E_{k+1}$  defined as  $re f(x_1, \dots, x_n) = f(x_1, \dots, x_n)$ , for all  $x_1, \dots, x_n \in E_k$ .  $I_A, A \subseteq E_k$ , is the set of all finitary functions on  $E_k$  that preserve the set  $A$ .

By definition,  $G_k$  is the clone of all finitary functions  $f$  on  $E_k$  that satisfy the following conditions:

1. For every  $g \in der f$ , if  $g \neq c_k$ , then  $g \in I_{\{0, \dots, k-1\}}$ .
2. If  $f(a_0, \dots, a_{n-1}) \neq k$ , for some  $a_i \in E_{k+1}$  and  $a_{i_0} = \dots = a_{i_j} = k$ , for  $0 \leq j \leq n-1$  and  $0 \leq i_0 \leq \dots \leq i_j \leq n-1$ , then the function

$$re f(a_0, \dots, a_{i_0-1}, x_1, a_{i_0+1}, \dots, a_{i_j-1}, x_j, a_{i_j+1}, \dots, a_{n-1})$$

is constant.

**4. Theorem.** Let  $F$  be an interpreted language with constants on  $E_k$  such that every function in  $F^{(1)}$  has a fixed point and  $F^{(1-1)} = st(k)$ . Then either  $F \subseteq G_k$  or  $F \subseteq Mon(\leq_k)$ .

**5. Theorem.** Any interpreted language  $F$  with constants on  $E_k$  such that every function in  $F^{(1)}$  has a fixed point and  $F^{(1-1)} = st(k)$  has the fixed-point property.

**6. Corollary.** The four-valued interpreted language that contains the constants and the operators  $\neg b, \neg_*, \wedge_{sw}, \downarrow_1^*, \downarrow_2^*, \downarrow_3$  has the fixed-point property.

## References:

[1] Melvin Fitting, Kleene's three-valued logics and their children, *Fundamenta Informaticae*, 20, pp. 113-131, 1994.

[2] Anil Gupta and Nuel Belnap, *The Revision Theory of Truth*, MIT Press, 1993.

[3] Albert Visser, Semantics and the Liar Paradox, in D. Gabbay and F. Guentner (eds.), *Handbook of Philosophical Logic*, vol. 4, Reidel, 1984, pp. 617-706.

# Naturalising Modal Contents

Manolo Martínez

Logos – Logic, Language and Cognition Research Group  
Departament de Lògica, Filosofia I Història de la Ciència  
Facultat de Filosofia (UB) C/ Montalegre, 6, 08001 Barcelona  
manolomartinez@ub.edu

## 1. The Problem.

It is a common assumption that teleosemantics, and other broadly causal-informational theories of content, are unable to account for Van Inwagen has called “the ordinary human powers of “modalization”” (Van Inwagen 1998, p. 73); i.e., our ability to entertain, and obtain justification to believe, contents involving modal properties. In this paper I aim at providing the foundations for a broadly causal-informational account of our sensitivity to modal properties even if, of course, such sensitivity in no way involves the causal powers of the uninstantiated.

## 2. A Toy Account of Content.

Throughout the paper I presuppose the following account of the attribution of contents to (selected-for, simple) mental states.

A mental mechanism, M, can be in one of only two possible states *-on or off*. We want to provide a content attribution of the kind “There is an F around” to M's being *on*. fact that M has been selected for is grounded on the co-occurrence of a cluster of properties in the environment in which M has thrived -e. g., in the infamous example of the frog and the flies, those properties include, among many others: *Being a black speck* and *Being nutritious for frogs*. Said cluster co-occurs because a certain homeostatic mechanism is in place -e.g., the mechanism of fly reproduction. Property cluster plus homeostatic mechanism individuate a *homeostatic property cluster* (HPC) in, roughly, Boyd's sense (cf., e. g., his 1991).

Finally, the proposal is that the content of M's being *on* is “There is a F around”, where F is the homeostatic property cluster that explains, in the way just sketched, M's being selected for (in the example, F is *fly*).

## 3. Contents Involving Probabilities.

In some cases, an agent may take advantage of the probabilistic relation existing among properties. The following variation on the frog/fly example is a case in point:

Not all flies are nutritious for frogs. In fact, only a 30% of them are (we will call these *goodflies*); the rest are harmless, but they report no advantage to the frogs who eat them. Three different mechanisms have mutated into existence in the brain of three frogs:

- Democritus has a mechanism M that makes him hunt whenever he sees a fly. His success rate in hunting is high
- Epicurus has a mechanism N that makes him hunt only when he sees a goodfly. His fitness-gains per hunt are bigger, then. On the other hand, his success rate is low: it takes time to tell goodflies from the rest of flies, and this allows them to discover the hunting frog and start fleeing from it.
- Empedocles has a slightly more complicated strategy. When he sees a fly he enters in a state we may call PREPARATION. PREPARATION is such that, once it is on, if

Empedocles goes on to hunt -because his mechanism N goes on-, his success rate will be a high. On the other hand, switching PREPARATION on has a low cost in resources.

It can be shown that, under a number of different assumptions about the benefits and costs of hunting, Empedocles's strategy is superior in the long run to Democritus's and Epicurus's. This, analogously to our toy theory of content attribution above, makes reasonable an attribution to the state in Empedocles's brain composed of PREPARATION and N with the content: "P(goodfly|fly) is X".

I then provide a general, naturalistic recipe for the attribution of probability-involving contents to mental states along these lines.

#### **4. From Probabilities to Possibilities.**

Using the ability to entertain contents involving probabilities as the main building block, I then go on to propose an analysis of our full-blown modal talk:

- It is possible that p if there is a time at which the probability of p is nonzero.

Some of the advantages of the proposal: it provides an explanation for our judgements of proximity among possible worlds -they are based on our predictions of the rate of change of the state of the world as a function of time-; it is thoroughly naturalistic -both our entertaining contents involving probabilities and contents involving the past can be given straightforward teleosemantic analyses-; it yields an analogue of the rationalist principle that conceivability is a guide to possibility -the very conditions for entertaining modal concepts are such that there is a connection between such entertaining and their being applicable, under specifiable circumstances.

#### **References**

- Boyd, R., (1991). 'Realism, Anti-Foundationalism and the Enthusiasm for Natural Kinds'. *Philosophical Studies*, 61:127-148
- Godfrey-Smith, P., (1996). *Complexity and the Function of Mind in Nature*. Cambridge University Press.
- Millikan, R., (1984). *Language, Thought and Other Biological Categories*. The MIT Press,
- Peacocke, C. (1999). *Being Known*. Oxford University Press.
- Van Inwagen, P. (1998) 'Modal Epistemology'. *Philosophical Studies*, 92:67-84.

## **Logic and Judgment: Frege versus Field<sup>1</sup>**

Concha Martínez-Vidal

Department of Logic and Moral Philosophy  
Universidade de Santiago de Compostela (Spain)

[mconcepcion.martinez@usc.es](mailto:mconcepcion.martinez@usc.es)

My purpose is to assess Frege's and Field's view about logical justification and its relation to judgment and rationality, where I take—following Travis—judgment to be “a posture towards the world, held rightly or wrongly (in one sense) solely in virtue of the world being as it is. A judgment fixes how the world would matter to its correctness; the world then does its mattering. [...] According to Frege's conception of judgment (Urteil) a posture may fail to be judging by failing to ask the world to matter in any substantive way. In the case of logic one might suspect the world to fail to matter”. My purpose will be to explore Frege's and Field's analysis of the problem to elaborate a comparative analysis on the kind of solution they advocate for the problem of logical justification. I will also try to determine whether Field's defense of a non-factualist view about justification is or is not independent from/consistent with his other philosophical tenets (physicalism, deflationism about truth among them).

I will start by introducing their views on logic and logical justification, signaling similarities and differences. One driving idea that seems to underlie both authors is that logic is to be relevant to rationality. In Frege's case this is obvious; but Field aims at defending that thesis too in the light of different results about logic that were not available in Frege's days.

My contention will be that similarities in their views about logic are to be explained in terms of their both belonging in the so called universalist conception (Heijenoort 1967) of logic and in the coincidences in their conception of truth. This is not to say that their views of logic or truth are completely coincident. They are actually divergent. One fundamental discrepancy between their conceptions of logic is that while Frege sustains there are logical entities, Field denies it. This conveys Field cannot accept that proof theory or model theory are adequate theories about logical facts because they use (or they are) mathematics (more on this below). Thus, he needs to clarify what he understands by 'logic' since logic is not to use mathematics either.

I will also explain their different views about the sources of the normative character of logic, about its a priori character, and about logical justification in terms of their sustaining radically different epistemological and ontological views and in terms of the way in which the amount of logical knowledge available allowed, in Frege's case, for some views that nowadays can no longer be sustained in his terms.

Their views about justification differ in that Field rejects there is such a thing as the kind of “justificatory fluid” Frege seems to assume there is. Clarifying what Field's non-factualism amounts to is a complicated issue. I understand that comparing his view with Frege's will help clarify his position. Field takes evaluations to be relative, not absolute (compare to his universalist conception of truth). The basic epistemological

---

<sup>1</sup> This research was supported by the FFI2009-08828 of the Spanish Ministry of Science and Innovation and FEDER (Cognitive Attitudes and Knowledge Justification) and by the Portuguese Government (“The Bounds of Judgment” a project of the MLAG Group of the Institute of Philosophy (Porto University) PTDC/FIL-F=L/109882/2009).

properties are not such properties as that of belief B being justified, but rather such relativized properties as that of belief B being justified relative to evidential system E (evidential system: a bunch of rules for determining under what conditions one is to believe various things; a belief is justified relative to an evidential system in certain circumstances if the rules license the belief under those circumstances). Now, to make relativism a believable doctrine, he has to show that all legitimate epistemological reasoning can be reconstructed with such relativized notions and without the absolute or unrelativized ones. The problem is, of course, whether we want to be able to say only that a given evidential system is better than another, or that a given evidential system is correct or best. Field contends evidential systems have lots of factual properties besides the properties of being adequate idealizations of the epistemic behavior of particular people. One factual property they have is reliability. An epistemological relativist presumably will want to bring his belief forming behavior into accordance with *some* reliable evidential system. Thus, to say that a belief is justified relative to *some* highly reliable evidential system is the highest epistemological praise. This is different from the absolutist view point (Frege's among others): one evidential system stands out from the rest as the true evidential system. Thus, for Field, even if reliabilism is one desirable property for an evidential system to have, there is no best evidential system to be discovered.

## References

- Field H. 1989, *Realism, Mathematics and Modality*, Basil Blackwell
- Field, H., 2009, "[Epistemology Without Metaphysics](#)", *Philosophical Studies*, 143, 249–290
- Field H., 2009b, "Logical Pluralism", *The Review of Symbolic Logic*, vol.2, no.2, 342-359.
- Field H., 2009c, "What is the Normative Role of Logic", *Aristotelian Society Supplementary Volume*, 83, Issue 1 , 251 – 268.
- Frege G., 1879/1967, *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*, Halle a. S.: Louis Nebert. Translated as *Concept Script, a formal language of pure thought modelled upon that of arithmetic*, by S. Bauer-Mengelberg in J. vanHeijenoort (ed.), *From Frege to Gödel: A Source Book in Mathematical Logic, 1879-1931*, Cambridge, MA: Harvard University Press, 1967.
- Frege G., 1884/1974, *Die Grundlagen der Arithmetik: eine logisch-mathematische Untersuchung über den Begriff der Zahl*, Breslau: W. Koebner. Translated as *The Foundations of Arithmetic: A logico-mathematical enquiry into the concept of number*, by J.L. Austin, Oxford: Blackwell, second revised edition, 1974.
- Frege G., 1893, *Grundgesetze der Arithmetik*, Jena: Verlag Hermann Pohle, Band I. Partial translation as *The Basic Laws of Arithmetic* by M. Furth, Berkeley: U. of California Press, 1964.
- Frege G., 1918/1956/1977, "Der Gedanke" in *Beiträge zur Philosophie des deutschen Idealismus*. Translated as "The Thought: A Logical Inquiry", *Mind*, V.65, 259, 289-311.
- Frege, G. 1967/84, *Collected Papers on Mathematics, Logic and Philosophy*, Edited by Brian McGuinness, Basil Blackwell: Oxford.
- Frege G., 1969/1979, *Posthumus Writings*, Hermes H., Kambartel, F, Kaulbach, F. (eds), , Basil Blackwell: Oxford.
- Heijenoort J. van, 1967, "Logic as Calculus and Logic as Language", *Synthese*, vol. 17, 324-330.

**Dependencia racional y autonomía epistémica:  
acerca del anti-individualismo y el papel de los expertos**

Juan V. Mayoral de Lucas

Departamento de Filosofía / Universidad de Zaragoza

Pedro Cerbuna, 12 /50009 Zaragoza

[jmayoral@unizar.es](mailto:jmayoral@unizar.es)

Las actuales tendencias anti-individualistas o comunitaristas en epistemología conceden un papel fundamental a la figura del experto como fuente de creencia racional. Uno de estos defensores del anti-individualismo, Sanford Goldberg, lo ha expuesto de manera clara mediante el siguiente ejemplo: “Sally no es científica. Depende de otros para lo que sabe acerca de los electrones, las cámaras de iones, el platino y las heliobacterias. Ellos le dicen cosas tales como que los electrones tienen espín, que las cámaras de iones detectan la radiación de ionización, que el punto de fusión del platino está por encima de los 1700 °C, y que el nicho ecológico de las heliobacterias [...] es el intestino ciego” (2009, p. 581).<sup>1</sup> Goldberg añade, con razón, que la dependencia epistémica que este ejemplo ilustra no es novedad. En las tendencias anti-individualistas arriba mencionadas se ha hecho énfasis en este punto para adentrarse en la naturaleza irreduciblemente social del conocimiento. Algunos autores (bien conocidos) de esta orientación incluyen a Tyler Burge, Elizabeth Fricker, John Hardwig, Martin Kusch, Frederick Schmitt y Michael Welbourne, entre otros.

Mientras que el anti-individualismo del que hablo nos ofrece una imagen plausible y sólida del fundamento de la creencia racional en un entorno social, dicha imagen se obtiene en él sólo a cierto precio. En particular, sus ideas acerca del *status* epistémico de los expertos sustituyen una imagen *igualitarista* de la distribución *a priori* de las capacidades epistémicas y de decisión racional en cualquier sociedad por algo más parecido a una representación *oligárquica*: dependemos irreduciblemente de los expertos y, en el ejercicio de nuestra responsabilidad epistémica, *debemos* someternos a su dictado. Así, en Fricker, por ejemplo, encontramos una afirmación como la siguiente: “[...] donde sé que otro [sujeto] es un experto en términos epistémicos en relación a mí y acerca de alguna cuestión, no es sólo para mí racionalmente tolerable, sino racionalmente *obligatorio*, aceptar su juicio en lugar del mío, en la medida en que tenga base suficiente para confiar en su sinceridad” (2006, pp. 242-243; cursivas mías). De modo semejante, Kusch señala que: “Creer a los expertos no sólo es un derecho epistémico; en muchos casos podría ser incluso una *obligación*” (2002, p. 47; cursivas mías). Podemos encontrar argumentos semejantes en, e.g., los textos bien conocidos de John Hardwig sobre esta cuestión (cf. esp. Hardwig 1985). El radical cambio de perspectiva acerca de la autonomía epistémica que este anti-individualismo impone merece como mínimo una revisión crítica detallada. En esta comunicación sólo iniciaré dicha revisión y para ello empezaré por tratar los argumentos de los tres filósofos que acabo de citar: Fricker, Hardwig y Kusch.

Mi doble tesis será la siguiente:

- (1) podemos evaluar al experto en términos epistémicos, pese a que
- (2) ciertos tipos de educación, y algunas habilidades cognitivas y prácticas,

---

<sup>1</sup> Todas las traducciones realizadas a continuación son mías.

hacen del experto una autoridad en términos epistémicos

Hardwig y Kusch, por ejemplo, eluden (1) sobre la base de su incompatibilidad con (2), mientras que yo sostengo que (1) y (2) son perfectamente compatibles —e incluso que *deben* serlo para que la propia perspectiva epistemológica de estos autores se sostenga— y que todo lo que necesitamos para combinarlas (i.e., (1) y (2)) es asumir que, con frecuencia, tomamos decisiones prácticas que son plenamente racionales. La propia Fricker, junto a Alvin I. Goldman (2001) y C. A. J. Coady (1992) ya ofrecen visos de esta compatibilidad, pero no la hacen explícita ni, por supuesto, contribuyen grandemente a su consolidación. En esta comunicación trataré de mostrar que (y cómo) puede lograrse.

En la §1 mostraré las perspectivas de Fricker, Hardwig y Kusch a este respecto como un único argumento que poder debatir. Propondré además algunas de las primeras dudas que a mi juicio suscita dicho argumento común. En la §2, pasaré a exponer los fundamentos de esa posición común y sugeriré lo que considero que es su punto débil. En la §3 me referiré a la tesis (pretendidamente) alternativa de Alvin Goldman (2001) y trataré de mostrar por qué creo que no es tal cosa. Finalizaré con dos secciones (§§4-5) en que esbozo mi propia perspectiva, ya antes apuntada. Dicha perspectiva se basa, curiosamente, en el trabajo de un anti-individualista como Sanford Goldberg. Emplearé su matiz del experto no sólo en términos *epistémicos* sino también *semánticos* para mostrar de qué manera las tesis (1) y (2) pueden mostrarse compatibles. La brevedad del sumario no me permite ampliar más mi argumento. Baste decir que dicho matiz de Goldberg ayuda a ver que es posible respetar dominios de conocimiento experto, a cuya autoridad recurrimos cuando (racionalmente) delegamos en ellos como autoridades epistémicas, a la par que esa concesión puede continuar siendo el resultado de una conclusión razonada (y, sobre todo, fruto del diálogo). Así no prescindiremos del igualitarismo de partida antes referido, aunque admitamos un alto grado de dependencia epistémica (y también semántica), que es resultado de una decisión racional y práctica.

Como dice Anthony Coady: “[...] el pensador autónomo no tiene por qué renunciar por completo a cierta medida de dependencia fundamental de la palabra de los demás, más bien debe utilizarla para lograr una actitud crítica verdadera y una perspectiva independiente viable. Todos necesitamos autonomía intelectual para alcanzar un grado factible de control sobre la creencia que adquirimos y para asegurar que nuestro pensamiento es adecuadamente sensible a nuestra historia cognitiva real y al medio intelectual presente”. (1992, p. 99; cursivas mías.) Mi contribución tratará de ampliar en lo posible este, ya con frecuencia citado, apunte de Coady.

## Referencias

- Coady, C. A. J. (1992). *Testimony: A Philosophical Study*. Oxford: Oxford University Press
- Fricker, E. (2006). “Testimony and Epistemic Autonomy”. En J. Lackey y E. Sosa (eds.), *The Epistemology of Testimony*, Oxford: Oxford University Press
- Goldberg, S. (2009). “Experts, Semantic and Epistemic”. *Noûs*, 43: 581-98
- Goldman, A. I. (2001). “Experts: Which Ones Should You Trust?”. *Philosophy and Phenomenological Research*, 63 (1): 85-109
- Hardwig, J. (1985). “Epistemic Dependence”. *Journal of Philosophy*, 82 (7): 335-49
- Kusch, M. (2002). *Knowledge by Agreement*. Oxford: Oxford University Press

## La regla $\delta$ , $\Pi'$ y E

José M. Méndez y Gemma Robles

Universidad de Salamanca. Edificio FES, Campus Unamuno, 37007, Salamanca.

Correo-e: sefus@usal.es; URL: <http://web.usal.es/sefus>

Dpto. de Hist. y Filo. de la CC, el Leng. y la Ed., Universidad de La Laguna. Facultad de Filosofía,  
Campus de Guajara, 38071, La Laguna, Tenerife.

Correo-e: gemmarobles@gmail.com; URL: <http://webpages.ull.es/users/grobles>

En 1956 Ackermann publicó su extraordinario artículo “Begründung einer strengen Implikation” [1]. Define en él Ackermann dos sistemas axiomáticos,  $\Pi'$  y  $\Pi''$ . El primero es una lógica carente de las paradojas de la implicación estricta; el segundo es una extensión del primero en la que, utilizando las ideas de Johansson (cf. [5]), Ackermann introduce en  $\Pi'$  el concepto de modalidad propio de  $S_4$ .

Como es sabido, los sistemas de Ackermann son el origen de la lógica de la relevancia (cf. [3], [4]). Una de las reglas de uno y otro sistema es la regla  $\delta$ , que reza:

“Si  $B$  y  $A \rightarrow (B \rightarrow C)$ , entonces  $A \rightarrow C$ ”.

En relación con esta regla Ackermann anota lo siguiente ([1], pp. 119-120):

“Für die Gültigkeit von ( $\delta$ ) im Gegensatz zu ( $\alpha$ ), ( $\beta$ ) und ( $\gamma$ ), ist wesentlich daß  $B$  eine logische Identität ist. Die Regel ( $\delta$ ) könnte also nicht in dieser form bestehen bleiben falls nicht logische Zusatzaxiome neben (1)-(15) auftreten”.

( $\alpha$ ) es *modus ponens*, ( $\beta$ ) la regla *adjunción*, y ( $\gamma$ ) el *silogismo disyuntivo*, y (1)-(15) los axiomas de  $\Pi'$ . En relación con ( $\alpha$ ), ( $\beta$ ) y ( $\gamma$ ), Ackermann comenta ([1], p.125):

“Die allein übrig bleibenden Regeln ( $\alpha$ ), ( $\beta$ ), ( $\gamma$ ) sind so gefaßt, daß sie auch bei Hinzufügung von nicht logischen Zusatzaxiomen gültig bleiben”.

Es, pues, claro que, para Ackermann, ( $\alpha$ ), ( $\beta$ ) y ( $\gamma$ ) son válidas como *reglas de inferencia*: pueden aplicarse a premisas cualesquiera. Pero  $\delta$  sólo es válida en la forma siguiente:

“Si  $B$  es demostrable y  $A \rightarrow (B \rightarrow C)$ , entonces  $A \rightarrow C$ ”.

Pues bien, en [6] se demuestra que  $\delta$  y la regla *aserción*

aser. “Si  $A$  es demostrable, entonces  $(A \rightarrow B) \rightarrow B$  es demostrable”

son equivalentes respecto de una lógica tan débil como la lógica  $B_+$  (“Basic Positive Logic”) de Routley y Meyer (cf. [7], [9]).

Es, por tanto, evidente que, mientras ( $\alpha$ ), ( $\beta$ ) y ( $\gamma$ ) son reglas de inferencia, aser es *regla de prueba*: sólo puede aplicarse a teoremas.

Anderson y Belnap definen su sistema E (“The Logic of Relevance and Necessity”) a partir de  $\Pi'$  eliminando ( $\gamma$ ) como regla primitiva y sustituyendo  $\delta$  por

$$[A \rightarrow (B \rightarrow C)] \rightarrow [LB \rightarrow (A \rightarrow C)]$$

ó

$$LB \rightarrow [[A \rightarrow (B \rightarrow C)] \rightarrow (A \rightarrow C)]$$

donde  $LA =_{df} (A \rightarrow A) \rightarrow A$  (cf. [2], p. 75).

Demuestran (de hecho ya lo habían demostrado en [2]) que cualquiera de estos dos axiomas puede jugar el papel de  $\delta$  en el *fragmento implicativo* de E. Aunque no recordamos haber leído la prueba ya sea en [3] o bien en [4], no sería difícil demostrar que tal es también el caso en el *fragmento positivo* de E. Ahora bien, ¿es también el caso en la *lógica E completa*? En [6] se demuestra que  $\delta$  no es una *regla derivada* de E.

¿Es quizá una *regla admisible*? Pensamos que ciertamente  $\delta$  es una regla admisible en E, pero no conocemos ninguna prueba de este hecho, y hasta que tal prueba no se presente, E y  $\Pi'$  *no son equivalentes* “as to theorems” como se afirma en [4], Cap. VIII.

Pero pasemos a la semántica. En [8] se afirma que se proporciona una semántica para E,  $\Pi'$  y  $\Pi''$ . Ahora bien, los autores de [8] anotan:

“Thus normal E.m.s and indeed E.m.s are characteristic structures for  $\Pi'$  and  $\Pi''$ , i.e., they verify all and only theorems: however, these structures validate neither the rule of necessity nor ( $\delta$ ). In fact, it is impossible to design the base T of E.m.s so that these rules are verified in T. Suppose that the non-theorem  $\bar{A} \vee t \rightarrow A$ , i.e.,  $A \supset \Box A$  does not belong to T. Since all theorems of  $\Pi$  belong to T,  $\bar{A} \vee A \in T$ . Hence  $A \in T$  or  $\bar{A} \in T$ . But if  $A \in T$  and necessitation is verified  $t \rightarrow A \in T$ , whence  $\bar{A} \vee t \rightarrow A \in T$ , contradicting assumptions”.

Más aún, reconocen en relación con los modelos definidos ([9], p. 407):

“Admittedly, the presence of P, and its apparent uneliminability, makes the semantics of E more cumbersome and less attractive than that of some of its relevant rivals such as R and T”.

El propósito de la presente comunicación es definir una semántica que, además de validar el silogismo disyuntivo, valida también las reglas  $\delta$  y “necessitation” como Ackermann hubiera querido. Más aún, en nuestra semántica el conjunto P al que se refieren los autores de [9] en la cita anterior es innecesario. Por tanto, no es eliminable porque no existe.

#### *Agradecimientos*

- Trabajo financiado por los proyectos de investigación FFI2008-05859/FISO y FFI2008-01205/FISO del Ministerio de Ciencia e Innovación de España.
- G. Robles está respaldada económicamente por el Programa Juan de la Cierva del Ministerio español de Ciencia e Innovación.

#### REFERENCIAS

- [1] Ackermann, W., “Begründung einer strengen Implikation”, *Journal of Symbolic Logic*, 21/2, 113-128, 1956.
- [2] Anderson, A. R., Belnap, N. D. Jr., “The pure calculus of entailment”, *Journal of Symbolic Logic* 27, 19-52, 1962.
- [3] Anderson, A. R., Belnap, N. D. Jr., *Entailment. The Logic of Relevance and Necessity*, vol I. (Princeton University Press), 1975.
- [4] Anderson, A. R., Belnap, N. D. Jr., Dunn, J. M., *Entailment. The Logic of Relevance and Necessity*, vol II. (Princeton University Press), 1992.
- [5] Johansson, I., “Der Minimalkalkül, ein reduzierter intuitionistischer Formalismus”, *Compositio Mathematica*, 4, 119-36, 1936.
- [6] Méndez, J. M., “A Routley-Meyer semantics for Ackermann's logics of 'strenghe Implikation””, *Logic and Logical Philosophy* (aceptado).
- [7] Routley, R., Meyer, R. K. “Semantics of Entailment III”, *Journal of Philosophical Logic*, 1, 192-208, 1972.
- [8] Routley, R. Meyer, R. K., Plumwood, V., Brady R. T., “Semantics of Entailment IV” (in R. Routley et al., *Relevant Logics and their Rivals*, vol. 1, Atascadero, CA, Ridgeview Publishing Co.), Appendix I, 1982.
- [9] Routley, R. Meyer, R. K., Plumwood, V., Brady R. T., *Relevant Logics and their Rivals*, vol. 1. (Atascadero, CA: Ridgeview Publishing Co.), 1982.

## Speaker Reference and Singular Thought

Andrei Moldovan

LOGOS – Universidad de Barcelona

Departamento de Filosofía y Lógica y Filosofía de la Ciencia

Universidad de Salamanca,

Edificio F.E.S. Campus Miguel de Unamuno

37007 Salamanca España

[mandreius@usal.es](mailto:mandreius@usal.es)

Following Donnellan (1966), many authors agree that definite descriptions and other expressions have both *attributive* and *referential* uses. To say that an expression has a referential use is not to make semantic claim about it, but it is to say that a speaker *uses* an expression to refer an audience to an object (cf. Bach 2005: 3). Many authors, among which, Donnellan (1966: 288, 1968: 205), Kripke (1977: 395-6), Reimer (1998: 90), Neale (1990: 65), Ludlow&Neale (1997: 528-9), describe the phenomenon of speaker reference in similar ways. Given that not only definite descriptions, but also indefinite descriptions, proper names, demonstratives, indexicals and quantifier phrases, can be used referentially in appropriate contexts, the definition should be generalized. According to the accounts present in the literature, a speaker uses an expression *e*, part of a sentence *s*, to refer to an object *b* if and only if:

- (1) the speaker believes of *b* that *e* applies truthfully to it;
- (2) the speaker uses *s* to convey a singular thought about *b* (or to convey a singular proposition about *b*, or to talk about *b*);

The question that I pursue in this essay concerns the relevance of the phenomenon of speaker reference to semantics. One methodological approach takes identifying and characterizing speaker reference as methodologically prior to characterizing *semantic* reference. The former should be the basis for building a theory of the literal meaning of an expression type, in virtue of which the tokens of that expression have certain semantic properties. This methodological order “push back a question about linguistic reference to a question about how mental states – in particular, referential intentions – can be directed towards objects” (Siegel 2002: 19) One reason for this strategy is that we, as speakers, tend to have stronger intuitions about what we mean, and about what we refer to with our expression, than about the properties of the expressions themselves.

In line with the above definition, understanding the phenomenon of speaker reference depends on understanding what exactly it is to say that someone believes *of b*, or *about b*, that it is in a certain way, or to say that someone has *a singular thought* about *b*, or *entertains a singular proposition* about *b*. However, not any conception of singular thought (or of entertaining a singular proposition) could be paired with what it takes for a definition of referential uses to be adequate to the phenomena it attempts to describe. That is because, as it has been pointed out (cf. Strawson 1971: 77-8, Evans 1980:306), some referential uses of certain expression are such that the speaker intends that the audience identify perceptually or in memory the object she refers to. The referential intention captured in (2) is then the intention to put the audience in some epistemic state with respect to *b*. Evans points out that if a speaker says ‘That man is F’ pointing at someone in the context of utterance, or says ‘That man we met last night is F’, “the hearer can understand the remark only if he perceives the man concerned” (*idem* p.305), and respectively, “unless he actually remembers the man” (Evans: 306). In other cases what is required is that the hearer have certain information from testimony that she needs in order to identify the referent, such as when a speaker says

'That mountaineer is F', thinking of a certain mountaineer that was in the news lately. Similarly, indexicals like 'here' and 'I' require identifying the place and speaker for successful communication (or understanding, as Evans puts it). Evans points out that "the phenomenon is much more widespread, including most normal uses of proper names" (Evans 1980: 306).

On the other hand, there are referential uses, including many uses of proper names, in which the speaker refers to an object but does not intend in any way that the speaker identify or recognize the referent in any way. According to several authors, including Kripke (1980: 91-7), Soames (2002: 70), Sainsbury (2007: 113), being competent with a proper name does not require *knowing who* the referent is, but only having the right intention when using the name. In Soames' words, in order to be a user of a name *n* of an object *o*, "One must have acquired a referential intention that determines *o* as the referent of *n*. ... this may be done... by picking up *n* from others who used it as a name of *o*, and intending use *n* to refer to the same thing they did" Soames (2002: 70). Therefore, identification of the referent is not needed for proper use of names, and need not be part of what the speaker intends to achieve by her utterance. This means that the same name may be used with the intention that the hearer identifies the referent in one context, and without the intention to achieve identification, in a different context. Moreover, even when the intention to identify the referent is present, what exactly this amounts to is different from one case to another, as Evans has shown. As Boer&Lycan (1986) argue, *knowing who* is contextually dependent: in one context knowing who Aristotle is may just amount to knowing that he was the teacher of Alexander, while in a different context one only knows who Aristotle is if one is able to recognize an image of him among many others.

The lesson of all this I take it to be that, due to the contextual dependency of *knowledge wh-*, there is no simple way to derive the semantics of referential expressions from the phenomenon of speaker reference. The reason of the contextual dependency of what it takes for the referential intention in (2) to be fulfilled is that identification of the referent is a form of world knowledge, that is, extra-linguistic knowledge. The descriptive material conveyed by an expression used referentially functions as a search mechanism for identifying the referent, but only together with contextual information, i.e. extra-linguistic knowledge. My tentative conclusion is that the methodological order mentioned above can hardly be successful.

## Genuino ‘saber-cómo’: una defensa del anti-intelectualismo

Jesús Navarro Reyes

Dpto. de Metafísica y Corrientes Actuales de la Filosofía, Ética y Filosofía Política  
Universidad de Sevilla, C/Camilo José Cela, s/n, 41018, Sevilla, España

[jnr@us.es](mailto:jnr@us.es)

En este trabajo defiendo la distinción entre saber-cómo y saber-qué, expuesta originalmente por Gilbert Ryle en (1945) y (1950). A tal fin, me enfrento al intento de refutación de Stanley y Williamson (2001) y, con menos detenimiento, a otras formas del llamado “intelectualismo”. Según este posicionamiento teórico, el saber-cómo es una forma específica de saber-qué, sin cualidades lógicas o epistémicas particulares. Distanciándose de otros esfuerzos recientes de refutación (como Williams (2007) o Sgaravatti y Zardini (2008)), la postura aquí defendida otorga mucho menos al intelectualismo de Stanley y Williamson que, en mi opinión, no habrían logrado confrontar los aspectos más relevantes del anti-intelectualismo de Ryle.

En la primera parte del trabajo, mi intención es hacer dos reproches al planteamiento de Stanley y Williamson; en la segunda, hacer una defensa positiva de la distinción de Ryle y apuntar su relevancia para algunos debates contemporáneos.

El primer reproche que dirijo a Stanley y Williamson es que confunden en su reconstrucción del argumento de Ryle cierta tesis del propio autor con otras que Ryle atribuye al intelectualismo a fin de reducirlo al absurdo. La imposibilidad de concebir una suerte de “intermediarios esquizofrénicos” entre la acción (evento meramente mecánico) y el pensamiento (puramente interno, mental, y no agencial) es lo que refuta el intelectualismo y no, como sostienen Stanley y Williamson, la idea de que, puesto que *todo* pensamiento es acción, cada acción requeriría a su vez de una sucesión indefinida de acciones.

En mi segundo reproche, sostengo que Stanley y Williamson no defienden suficientemente la distinción que proponen entre atribución de habilidades y atribución de saber-cómo. En el planteamiento de Ryle ambos tipos de atribución son idénticos, y aduzco diferentes razones por las que es preferible mantener esa identidad. Principalmente, indico que el uso de la expresión “saber-cómo” es equívoca entre atribuciones efectivas de habilidades y atribuciones de saber-qué acerca de modos de acción (“ways” en la terminología de Stanley y Williamson). Sólo en caso de referirnos al primer uso estamos hablando de una atribución genuina de saber-cómo, que es precisamente atribución de *habilidad* al agente (ver Rosefeldt (2004)).

A partir de ahí, en la segunda parte del artículo es mi intención apuntar tres implicaciones del planteamiento anti-intelectualista que me parecen especialmente relevantes en el contexto actual. En primer lugar, el modo en que la atribución de habilidades excede las fronteras de lo tradicionalmente concebido como *mental*, extendiéndose no sólo al cuerpo del agente, sino incluso a los instrumentos de los que ha de hacer uso y a la materia en que ha de aplicarlos. Se trata por tanto de un principio central para la cognición situada y la hipótesis de la mente extensa.

En segundo lugar apunto una crítica a Ryle, señalando que parece más adecuado hablar de “invención” que de “descubrimiento” en los casos de adquisición de nuevos

métodos o prácticas. Entre otras razones, aduzco que, del mismo modo que el verbo “saber” se comporta de modo distinto según introduzca una cláusula proposicional o un infinitivo, el verbo “inventar” tiene un comportamiento manifiestamente distinto en “inventar-qué” e “inventar-cómo”. La epistemología del saber-cómo parece ser peculiarmente ajena a la noción de verdad, y no ha de extrañar por ello que, como ha sido apuntado en distintas ocasiones, la casuística de este tipo de conocimiento sea ajena a los llamados “casos de Gettier” (Gettier 1963).

En tercer y último lugar, pretendo señalar la importancia del planteamiento de Ryle para el debate actual acerca de las acciones mentales (ver O’Brien y Soteriou (2009)). Sostengo que una buena parte de la teoría de la acción de los últimos años ha intentado barrer debajo de la alfombra de lo *mental* las dificultades encontradas a la hora de analizar la conducta en términos agenciales (ver Mele (2003) y, para una visión crítica de ese planteamiento, Moya (1990)). La distinción de Ryle fuerza a abandonar este intento desafortunado.

### **Bibliografía:**

- Gettier, E.L (1963): “Is Justified True Belief Knowledge?” *Analysis* 23, 121-123.
- Lihoreau, F. (2008): “Knowledge-How and Ability”. *Grazer Philosophische Studien* 77, 263-305.
- McDowell, J. (1988). “Criteria, Defeasibility and Knowledge”. En J. Dancy, ed. *Perceptual knowledge*. Oxford: Oxford University Press, 209-19.
- Mele, A. (2003): *Motivation and Agency*. New York: Oxford University Press.
- Moya, C.J. (1990): *The Philosophy of Action: an Introduction*. Oxford: Polity-Blackwell.
- O’Brien, L. & Soteriou, M., eds. (2009). *Mental Actions*. Oxford: Oxford University Press.
- Rosefeldt, T. (2004): “Is Knowing-how Simply a Case of Knowing-that?” *Philosophical Investigations* 27:4, 370-379.
- Ryle, G. (1945): “Knowing How and Knowing That”. *Proceedings of the Aristotelian Society* 46, 1-16.
- Ryle, G. (1950): *The Concept of Mind*. Londres, Hutchinson.
- Sgaravatti, D. & Zardini, E. (2008). “Knowing How to Establish Intellectualism”. *Grazer Philosophische Studien* 77, 217-261.
- Stanley, J. & Williamson, T. (2001): “Knowing How”. *The Journal of Philosophy* 98:8, 411-444.
- Williams, J.N. (2007). “Propositional knowledge and know-how”. *Synthese* 165:1, 107-125.

## Essential demonstratives

Carlo Penco [penco@unige.it](mailto:penco@unige.it)

The paper aims to give a new classification of indexicals based on the relevance of complex demonstratives. By “demonstration” we mean any action or gesture which makes some element of the context salient or relevant. Sometimes it is said that the uttering of words is a kind of gesture, if nothing else is available. So why don’t we consider the uttering of “I” a gesture? This suggestion will guide us in the study of the relation between gestures and linguistic expressions. The best way to begin is with an analysis of the proper treatment of complex demonstratives, where the link between language and gestures is totally explicit. If the case of complex demonstratives could be generalized to all indexicals, this would better explain the peculiarity of the semantic category of indexicals as such. In fact, if we take context dependence as a phenomenon which is inclusive of the beliefs of the speakers, it is difficult to contrast indexicals with other context dependent expressions. A possible way to make a reasonable and sharp distinction between indexicals and other context dependent expressions is given by the connection with gestures; we may always express and disambiguate an indexical with an accompanying gesture, while we cannot do the same for other context dependent expressions like “enemies” and “friends”. Those context-dependent expressions cannot be disambiguated by pointing at them; you already need the concept of “enemy” and “friend” in order to use a gesture for referring to friends and enemies; but no *gesture* can *explain the concept* of enemy, even though we may point to what we believe to be enemies. This short remark contains the first suggestion of the paper: taking demonstration as central to indexicals is a clear means of creating a divide between indexicals and other context-dependent expressions.

I therefore propose a mental experiment where pure indexicals are treated as complex demonstratives. Let us suppose that to refer to the speaker in the context, instead of using “I” we used the complex indexical “this speaker” pointing towards ourselves. Similarly, to refer to the place of the context we might point to the place saying “in this place”; to refer to the time of the context we might say “in this time” or “in this day” (in latin “hodie” is exactly “in this day”). We normally use “this” to refer to periods of time: this week, this year, and so on. To indicate the relevant male/female of the context the speaker might say “this man/woman”, and so on.

Why do we pass from complex demonstratives to pure indexicals? Well, in case we are unable to see the demonstration of “this speaker”, an expression like “I” is certainly more suitable, especially when we recognize the voice of the speaker (here the tone of voice may be considered also a demonstratum in some kind of restricted linguistic community). Ambiguity may arise also in the case of “this time” or “this place”, but these ambiguities are common to “now” and “here”. If the indexical “I” may be used to avoid ambiguity of the complex demonstrative “this speaker”, a demonstrative however *may* always accompany any indexical in order either to stress (with “I” it is possible to point rhetorically to my heart or to my nose, depending on the culture) or to disambiguate: take the case of “he”/“she” or “him”/“her”, which may be accompanied by a gesture or a glance or the position of the body (possibly not a pointing one, which is normally not polite, although in some cases it can be used

without any problem). I don't mean that indexicals are *reducible* to demonstratives, but that they are entrenched with possible and actual demonstrations and may also be classified according to the distance from a possible demonstration.

The case of pure demonstratives ("this", "that") seems to lead us in the opposite direction: instead of having an explicit expression of the content, with pure demonstratives we have only the gesture, without any content. Is that true? It might be claimed that *any* use of a pure demonstrative is an attempt to select a sort of thing. Any occurrence of "this" or "that" triggers a possible answer to a "what?" question. A "what?" question is complex: it requires on the one hand a specific reference to a particular item, but on the other hand it requires a reference of a kind or sort of item. This is one of the standard lessons of Wittgenstein's *Philosophical Investigation* (but also of Bühler's theories): you cannot understand an ostensive definition of the kind "this is an F" unless you have already learned a lot in language. But what is it necessary to learn before understanding ostensive definitions? Well, probably basic differences among speech acts and basic categories. You cannot ask which color this is if you don't have any idea of what *a color* is. Your categorization might also be very confused (you may put very different things into a single category); still it is something on which you need to ground your question "what is this?" when pointing to a physical object, to a shape, to a color, to an event.

From these remarks a second suggestion follows: the use of demonstrations is something linked to a general ability to categorize *sorts* of things. This aspect is prior to the development of pure indexicals and may help to build a classification of indexicals depending on the relations between (i) whether the demonstration is explicit or implicit and (ii) whether the sortal connected with the demonstration is explicit or implicit. The classification I propose is not incompatible with the standard one described by John Perry, but has the advantage of offering a unified method of classifying indexicals and of separating indexicals from other context sensitive expressions.

Perry J. 2006, "Using Indexicals", in M. Devitt, R. Hanley, *The Blackwell Guide to the Philosophy of Language*, Blackwell (314-334).  
Textor M. 2007, "Frege's Theory of Hybrid Proper Names Developed and Defended", *Mind*, Vol. 116 (947-981).

## **Mapas y modelos. La representación científica de la realidad.**

Inmaculada Perdomo Reyes

Universidad de La Laguna

Tenerife, España

mperdomo@ull.es

El propósito de esta comunicación es mostrar cómo el modelo o analogía del mapa incorpora muchas de las características representacionales identificadas en la relación teoría-mundo, o dicho de otra forma, nos ayuda a entender cómo la ciencia representa el mundo, para destacar, a continuación, que los elementos interpretativos y nuestras necesidades intelectuales y pragmáticas juegan también un rol importante en nuestras construcciones modelo-teóricas del mundo. Estos elementos son introducidos en el debate desde la perspectiva empirista actual.

Lo más evidente, a primera vista, es que no existe un “mapa universal”, y que no tiene sentido preguntar si un mapa es verdadero o falso. Aunque sí puede definirse como empíricamente adecuado y lo mismo puede decirse de los modelos en ciencia. Las virtudes representacionales de los mapas son diferentes. Un mapa es más o menos preciso, más o menos detallado, o diseñado a mayor o menor escala. Además requiere la puesta en práctica de muchas convenciones humanas diferentes, tanto para su construcción como para su uso, y sin la comprensión y adecuada interpretación de tales convenciones no es más que una serie de líneas en un papel. Esta idea avanzada por R. Giere es nuestro punto de partida, que será contrastado también con las ideas de P. Kitcher, Bas van Fraassen y H. Longino, entre otros, para abordar las principales características y valores implicados en la construcción de los modelos en la ciencia y la tecnología en tanto prácticas representativas e interpretativas.

Desde esta perspectiva, en la que debemos considerar a la actividad científica misma más que a la ciencia como un cuerpo abstracto de conocimiento aceptado, lo más importante es destacar que tratamos con actividades intencionales, contextuales e históricamente condicionadas. Los usos de la teoría para explicar, las aplicaciones a la técnica, la interpretación de los datos generados por instrumentos científicos diseñados a tal fin y la construcción de modelos son actividades realizadas por las comunidades científicas y que requieren de una “localización” de los sujetos con respecto al cuerpo de conocimiento o información puesta en juego. Continuando con la analogía del mapa, la característica más importante de la construcción modelo teórica, no es su función representativa, con todos los matices que queramos incluir, sino el hecho de que constituyen instrumentos útiles para guiarnos o intervenir en el mundo, y en los que la interpretación es central. El mejor mapa, como dice H. Longino, es aquel que permite a sus usuarios conseguir sus metas. En esta comunicación queremos ahondar en esta perspectiva y argumentar a favor de una visión empirista y contextual de la práctica científica.

Pero los mapas son mapas acerca de algo, y a pesar de su falibilidad, limitaciones y dependencia de los elementos interpretativos y convenciones, esto es lo que debe destacarse a juicio de los autores comprometidos con una visión realista: el hecho de que este tipo de actividad proporciona un acceso genuino al conocimiento de ciertos aspectos de la realidad. Los modelos que construyen los científicos, sostiene R. Giere,

representan varios aspectos del mundo, aunque la relación ya no es definida en términos de verdad sino en términos de similitud en ciertos aspectos y en cierto grado, igual que en el caso de los mapas. Por su parte, P. Kitcher distingue entre el contenido propuesto (*intended content*) y las convenciones (*reading conventions*) en las que los elementos interpretativos entran en juego. En cualquier caso, si el mapa representa ciertas relaciones espacio-temporales a través de diferentes ítems, entonces existe una contrapartida de esas relaciones en el mundo.

Ahora bien, otro punto de vista desde el cual analizar o enfatizar otras características del modelo del mapa si confiamos en que es un buen ejemplo o analogía para explorar la forma en que la ciencia representa el mundo, es atender al acto mismo del uso del mapa. El poder representacional de un mapa puede ser atestiguado por cualquier persona que haya usado un mapa para guiarse en un territorio no familiar, pero, de hecho, para usarlo correctamente necesitamos información adicional que no está contenida en el mapa. Los mapas no le dicen al usuario dónde está, para localizarse en el mapa, e incluso cuando aquellos contienen la flecha indicando “usted está aquí”, el acto de la auto-localización con respecto a la flecha no es algo que el mapa delimite, como enfatiza van Fraassen. Tal acto, además no tiene que ver o no es deducible de su grado de precisión, ni puede ser identificado con el contenido o la creencia de que un mapa o modelo representa el mundo. Este aspecto no tiene cabida, o no pertenece al ámbito de la semántica sino al de la pragmática. Como Longino enfatizó en su obra de 1990, lo que constituye nuestro mundo no es algo dado, sino un producto de la interacción entre la realidad externa material y nuestras propias necesidades intelectuales y pragmáticas.

El empirismo actual, proporciona, a nuestro juicio, y así pretende ser defendido en esta comunicación, un análisis adecuado de la ciencia entendida como actividad constructora de modelos que permiten ofrecer representaciones e interpretaciones del mundo; un análisis adecuado del sujeto investigador y de la dinámica de las comunidades científicas, así como de los procesos sociales que autorizan o legitiman tales interpretaciones como conocimiento científico conforme o adecuado empíricamente.

#### Referencias bibliográficas:

- R. Giere, (1999), *Science without Laws*, University of Chicago Press.  
\_\_\_\_\_ (2006), *Scientific Perspectivism*. University of Chicago Press.  
P. Kitcher, (2001), *Science, Truth and Democracy*, Oxford University Press.  
H. Longino, (1990), *Science as Social Knowledge*. Princeton University Press.  
\_\_\_\_\_ (2001), *The fate of Knowledge*, Princeton University Press.  
B.C. Van Fraassen, (2002), *The empirical stance*, Yale University Press.  
\_\_\_\_\_ (2008), *Scientific Representation: Paradoxes of Perspective*. Oxford University Press.

## Escepticismo, normatividad y lo ordinario: Finkelstein y Cavell sobre seguir una regla

David Pérez Chico<sup>1</sup>

[dcperez@unizar.es](mailto:dcperez@unizar.es)

Universidad de Zaragoza

1. Las consideraciones sobre la cuestión de seguir una regla forman parte en la obra del último Wittgenstein de unas reflexiones más generales sobre el significado y la comprensión. No obstante, no cabe duda de que, por sí mismas, constituyen una de las cuestiones fundamentales del pensamiento wittgensteiniano, tanto para interpretar dicha filosofía, como por las implicaciones que dichas consideraciones tienen en otras cuestiones filosóficas, en especial en todas aquellas en las que la normatividad desempeña un papel importante.

2. El asunto de seguir una regla en Wittgenstein ha sido reconsiderado de manera sumamente interesante y sugerente por David Finkelstein en *Expression and the Inner (La expresión y lo interno*, [trad. de Lino San Juan], KRK, 2010). Con todo, el seguimiento de reglas no es el asunto principal del libro de Finkelstein; al contrario, su objetivo es proponer una concepción sobre la autoridad de la primera persona y la auto-ascripción de estados mentales estrechamente relacionada con lo que Wittgenstein afirma acerca de la expresión y los estados internos. No obstante, la concepción de Finkelstein se erige sobre una lectura previa de las observaciones wittgensteinianas sobre seguir una regla a partir de las lecturas de Crispin Wright, Saul Kripke y John McDowell.

Finkelstein argumenta de manera muy notable que los tres interpretan erróneamente lo que Wittgenstein afirma en IF §201. Y la razón, según nuestro autor, es que el problema no es que el contenido de una regla necesite ser interpretado, estipulado o justificado, sino más bien que la supuesta paradoja denunciada por Wittgenstein hunde sus raíces en una amenaza mayor: la que establece la existencia de un vacío entre una regla y su aplicación. Así, la lectura de Wright intentaría salvar el vacío a base de estipulaciones sobre cuál es la aplicación correcta; la de McDowell con la necesidad de justificar epistémicamente dicha aplicación; y la de Kripke aprecia la existencia del vacío, pero sostiene que cualquier puente que podamos tender para salvarlo se sostendrá sobre unos cimientos precarios.

Finkelstein, por su parte, sostiene que Wittgenstein no albergaba la intención de tender puente alguno, sino más bien trataba de convencernos de que no existe ningún vacío entre una regla y su aplicación. Que la aparente necesidad de tender puentes entre ellos responde únicamente a la creencia de que una regla es algo inanimado (como una ristra de manchas de tinta). Que los casos en los que parece abrirse un vacío son excepcionales, especiales o imaginados, por ejemplo, por los filósofos.

---

<sup>1</sup> Este trabajo forma parte del Proyecto de Investigación del Programa Nacional *Puntos de vista. Una investigación filosófica* (Ref: FF2008-01205).

3. Kripke es, opinablemente, el principal responsable del interés mostrado hacia el asunto de seguir una regla durante los últimos 28 años. De su lectura nos interesará la importancia que otorga al escepticismo en las reflexiones wittgenstenianas. Una importancia que desgraciadamente desaparece en Finkelstein. Desgraciadamente porque en otros aspectos su lectura me parece más próxima al núcleo de los pensamientos wittgenstenianos y, sin embargo, defenderé que presenta importantes limitaciones en algunos aspectos cruciales para captar el verdadero alcance del pensamiento wittgensteniano sobre seguir una regla. En especial respecto a la insistencia wittgensteniana en el carácter público de nuestro lenguaje. Es probable que lo que Finkelstein comenta acerca del “espacio público de la vida animada” sea contrario a mi opinión, pero intentaré mostrar que si queremos que la solución desarrollada en las secciones §§201-242 sea relevante para los problemas examinados por Finkelstein, lo que Wittgenstein afirma sobre las prácticas, las costumbres, el acuerdo en juicios y en formas de vida, lo dado, y el suelo de roca dura donde todas las justificaciones deben finalizar, tiene que ser tenido en cuenta más de lo que lo es en *Expression and the Inner*, en especial a la luz de una cierta lectura que no niegue la sensibilidad de Wittgenstein hacia la amenaza escéptica. Sospecho que, como sostiene Stanley Cavell, *mantener una relación con la posibilidad del escepticismo es un aspecto interno de la filosofía de Wittgenstein*.

4. Una primera aproximación a nuestro tema nos acerca al que considero que es el segundo error de la lectura kripkeana (el primero es que base su lectura en una paradoja que Wittgenstein descarta casi tan pronto como la enuncia): Kripke insista en calificar como escéptica la solución que atribuye a Wittgenstein después de haber rechazado la concepción que vincula el significado de una expresión con sus condiciones de verdad. Mantengo la opinión de que este error es aún más extraño a Wittgenstein, pues una de las enseñanzas que obtengo de la sección §201 es que la relación normativa entre una regla y su aplicación en particular, pero también nuestra relación con el mundo en general, podría haber sido de otra manera. Los malentendidos y las interpretaciones erróneas constituyen una amenaza constante en la filosofía de Wittgenstein, pero esto no significa que las cosas no sean como pensamos que son, o que no tengamos derecho a actuar como lo hacemos, o a atribuir a nuestras palabras el significado que tienen.

El principal problema no es, como mantiene Finkelstein, si Wittgenstein reconoce o no la existencia de vacíos, sino más bien qué tipo de puentes habría que tender para salvarlos (una clase de puentes que, a pesar de la amenaza constante del escepticismo, sea capaz de explicar la regularidad de nuestras existencias). Y no sólo nos hace ver Wittgenstein que ni las interpretaciones, ni un consenso comunitario servirían para salvar las posibles brechas, sino que al definir las reglas como prácticas, y al fundamentar el seguimiento de reglas y cualquier juicio sobre dicho seguimiento en un acuerdo más básico en juicios, los puentes que contemplaría Wittgenstein no son de los que unen dos extremos irreconciliables a la espera de ser cruzados (valga aquí como metáfora lo que, en *El libro de Manuel*, dice Cortázar: “un puente es un hombre cruzando un puente”).

5. El comentario de la lectura que hace Stanley Cavell de la cuestión de seguir una regla en Wittgenstein nos servirá para ilustrar nuestras conclusiones.

## **Acerca del tipo de conocimiento que María tiene en el momento de su liberación.**

Diana I. Pérez

Universidad de Buenos Aires- CONICET

El famoso "argumento del conocimiento" fue propuesto por Frank Jackson (1982, 1986) para probar la falsedad del fisicalismo. El argumento es bien conocido: propone la existencia de una neurocientífica que *sabe* todo lo que hay que saber del mundo físico, particularmente de la visión de color, pero que estuvo toda su vida confinada en un cuarto blanco-y-negro. Si se acepta que al salir de la habitación y ver algo rojo por primera vez, la neurocientífica, María, aprende algo nuevo, a saber "cómo es ver rojo", entonces estamos implícitamente aceptando que ahora *sabe algo que antes no sabía*, y por tanto que no sabía antes de salir *todo* lo relacionado con la visión de color. Pero si sabe algo nuevo, y ya sabía todo lo físico, esto nuevo no puede sino ser no-físico, por lo tanto el fisicalismo es falso.

Este argumento dio lugar a una extensa bibliografía filosófica en los últimos 30 años. En este trabajo voy a centrarme en un aspecto de este argumento, que es el relacionado con la idea de que hay un tipo peculiar de *conocimiento* que María adquiere al salir (la centralidad de este punto es indiscutible, dado el nombre con el que se ha popularizado el argumento). Entre las múltiples respuestas fisicalistas proporcionadas a este argumento, hay una gran variedad que se centra en la cuestión del *tipo de conocimiento* que María adquiere.

(1) En efecto, la lectura más directa del argumento parece suponer que la conclusión es que María conoce un hecho no físico al salir, y que por lo tanto el conocimiento de la propia experiencia subjetiva es un tipo más de *conocimiento de hechos, o conocimiento proposicional*. Esta idea, por lo demás, ha dado lugar a una extensa literatura alrededor de los peculiares constituyentes de este contenido proposicional de nuestra experiencia subjetiva: a saber los así llamado "conceptos fenoménicos" (Loar 1990, Tye 2000, 2003, Levin 2007, Papineau 2002, 2007 entre muchos otros) . Uno de los objetivos del trabajo es evaluar las diferentes propuestas alrededor de los conceptos fenoménicos, mostrando que ninguna de las teorías existentes parece satisfactoria.

(2) Una segunda opción fisicalista explorada ha sido la propuesta por Lewis 1988, y Nemirow 1990, consistente en sostener que el tipo de conocimiento que adquiere María al salir no es un conocimiento proposicional sino un "*know how*", esto es un nuevo conjunto de habilidades (cognitivas). Esta propuesta también ha sido extensamente discutida, tanto porque no parece proporcionar una explicitación adecuada de lo que le ocurre a María al salir (por ej porque las habilidades requieren algún tipo de entrenamiento, y por lo tanto no parecen ser el tipo de cosa que María pueda adquirir instantáneamente al salir, tal como se presupone en el argumento de Jackson), como porque bajo alguna lectura de la naturaleza de los conceptos, esto es entendiéndolos como habilidades cognitivas, no parece haber una diferencia sustantiva con la posición anterior.

(3) Una tercera propuesta, recientemente formulada por Tye (2009 cap. 5), pretende dar cuenta de la peculiar situación de María recuperando la idea de que la experiencia perceptiva incluye un cierto *conocimiento directo (acquaintance)* que no pudo tener María antes de salir, y que es la clave para entender su situación epistémica al salir. Conee es el que defiende que se trata de conocimiento por acquaintance.

Otras posibilidades, tal vez no exploradas en detalle pero mencionadas en la literatura, son las siguientes.

(4) Se puede suponer que el conocimiento de nuestra propia experiencia subjetiva es un tipo de *conocimiento de objetos* (como conocer Atenas o al Papa) (Tye 2009, cap 5), algo que está conectado con el conocimiento directo mencionado arriba, pero que no supone la existencia de datos de los sentidos como aquellos particulares con los que se establece tal conexión epistémica. La diferencia entre esta propuesta y la anterior parece estar en la naturaleza del objeto conocido.

(5) También se puede considerar que lo que adquiere María es un caso de *conocimiento proposicional*, pero sin involucrar conceptos peculiares (i.e. conceptos fenoménicos), sino sólo los *conceptos psicológicos ordinarios* (Pérez 2008, 2009), es decir asimilando los conceptos involucrados en este tipo de conocimiento a otros conceptos deferenciales (Tye 2009, cap 3) o de clases naturales (Pérez 2005)

Todas estas opciones teóricas podrían ser relevantes para dar cuenta de en qué consiste "saber cómo es ver rojo". Pero todas las opciones (argumentaré) presentan dificultades. Podría entonces, explorarse la idea de que no es correcto describir la situación de María al salir afirmando que adquiere un nuevo *conocimiento*, después de todo, sino describiendo su situación en términos de la ocurrencia de algún otro tipo de estados psicológicos, por ej de percatación (*awareness*) o simplemente un sentir (*feel*) lo que es experimentar rojo. En el caso de la percatación se trata de un estado con contenido conceptual, en el caso del mero sentir no. En la última parte del trabajo exploraré estas dos alternativas, esto es trataré de mostrar que no hay un nuevo conocimiento que efectivamente María adquiere en el acto mismo de su liberación, tal como se ha supuesto a partir del argumento de Jackson, sino que más bien lo que ocurre es otra cosa, que *da lugar a* la posibilidad de un nuevo conocimiento. El contraste entre el caso imaginario de María y el caso real de aprendizaje de los conceptos con los que aprendemos a conceptualizar nuestra experiencia subjetiva será crucial para probar mi sospecha.

### **Referencias bibliográficas.**

- Jackson, F. (1982) "Epiphenomenal qualia", *Philosophical Quarterly*, XXXII, 127: 127-136
- Jackson, F. (1986) "What Mary Didn't Know" *Journal of Philosophy*, 83:5, pp. 291-5.
- Levin, J. (2007) "What is a phenomenal concept" en Alter, T and Walter, S (2007) *Phenomenal concepts and phenomenal knowledge*, OUP.
- Lewis, D. (1988) "What experience teaches" reprinted in Block, B, Flanagan, O and Guzeldere, G (2002) *The nature of consciousness*. MIT Press.
- Loar, B. (1990) "Phenomenal states" reprinted in Block, N., Flanagan, O, and Guzeldere, G. *The Nature of Consciousness*, MIT Press.
- Papineau, D. (2002) *Thinking about Consciousness*. OUP.
- Papineau, D. (2007) "Phenomenal and perceptual concepts" en Alter y Walter 2007.
- Pérez, D. (2004) "Mental concepts as natural kind concepts" CJP.
- Pérez, D. (2008) "Why should our mind-reading abilities be involved in the explanation of phenomenal consciousness?" *Análisis Filosófico*, vol. XXVIII, no. 1.
- Tye, M. (2003) "A Theory of Phenomenal Concepts", en O'Hear, Anthony (ed.) *Minds and Persons: Royal Institute of Philosophy Supplement 53*, Cambridge University Press, Cambridge, pp. 91-105.
- Tye, M. (2009) *Consciousness Revisited. Materialism without phenomenal concepts*. MIT Press, Cambridge.

## Defensa de una versión del argumento trascendental de Putnam contra la posibilidad epistémica de que seamos cerebros en una cubeta\*

Manuel Pérez Otero

Departamento de Lógica, Historia y Filosofía de la Ciencia. (Facultad de Filosofía. Universidad de Barcelona.) / LOGOS (Research Group in Logic, Language and Cognition).

C / Montalegre 6. 08001 Barcelona. E-mail: [perez.otero@ub.edu](mailto:perez.otero@ub.edu)

H. Putnam (1981) propuso un argumento mediante el cual pretendía demostrar que no somos *cerebros en una cubeta* (esquemáticamente: cerebros desprovistos de cuerpos, que experimentarían sensaciones –provocadas artificialmente– subjetivamente indistinguibles de las sensaciones que experimenta un sujeto normal). Incluso muchas personas que pudieran tener inclinación a simpatizar no sólo con la conclusión sino también con la estrategia argumentativa de Putnam (de carácter trascendental) reconocen al menos cierto aire paradójico en el razonamiento que Putnam presentaba. Es muy difícil ofrecer una interpretación literal de su texto que evite la paradoja. No será ese mi propósito. En su lugar, presentaré y defenderé una versión algo diferente de su argumento, cuya diferencia principal más relevante es –precisamente– que en esta versión no hay riesgo de caer en una paradoja. Es un argumento inspirado en el de Putnam; podría razonablemente considerarse que algo así estaba en su mente. Pero la cuestión exegética sobre si cabe atribuirle o no esta versión no es mi tema principal aquí.

Mi interpretación del argumento putnamiano no es completamente original. Está directamente inspirada por una discusión en que Falvey y Owens (1994) consideran y rechazan un argumento muy similar. En dos aspectos se separa mi posición de la suya, siendo el segundo más importante:

(a) Propongo añadir una premisa adicional al argumento que Falvey y Owens discuten, sin la cual dicho argumento es defectuoso.

(b) Falvey y Owens contemplan la posibilidad de complementar el argumento con una premisa de ese tipo. Rechazan el resultado alegando –entre otras razones– que entonces el razonamiento cometería petición de principio. Mi objetivo fundamental es explicar por qué no es problemático el argumento con esa premisa adicional. No obstante, la discusión permitirá constatar el alcance real de la argumentación putnamiana (conforme a la interpretación que aquí propondré).

El argumento putnamiano que respaldo puede formularse del siguiente modo:

- (A) P1: Estoy considerando el pensamiento de que no soy un cerebro en una cubeta  
P2: Es verdadera una teoría mínimamente causal y externista sobre el significado  
P3: Si  $x$  fuera un cerebro en una cubeta, entonces  $x$  no podría considerar el pensamiento de que no es un cerebro en una cubeta

-----  
C: Por tanto: No soy un cerebro en una cubeta

Ciertos supuestos usuales sobre la autoridad especial de la primera persona en relación con los contenidos de los propios pensamientos respaldan suficientemente la premisa P1, de forma que un sujeto que cree que no es un cerebro en una cubeta tiene conocimiento de lo que P1 establece. Respecto a P3, esa proposición derivaría de P2. Por limitaciones de espacio, no especificaremos cómo podría derivarse P3 a partir de P2.

---

\* Financiación: Programa CONSOLIDER-INGENIO 2010, “Perspectival Thoughts and Facts” (CSD2009-00056), MICINN. Proyecto “Discriminabilidad: representación, creencia y escepticismo” (FFI2008-06164-C02-01), MICINN. Grupo de investigación consolidado y financiado LOGOS (2009 SGR 1077), DIUE.

Falvey y Owens probablemente aceptarían que P2 implica P3, pues mencionan una relación similar entre una tesis algo diferente a P2, más específica, y P3. Pero objetan que dicha tesis es una proposición empírica (como también lo es P2) que sería falsa si fuéramos cerebros en una cubeta. Por ello, emplear dicha tesis para justificar P3 en el contexto de una argumentación cuya conclusión es C (“no soy un cerebro en una cubeta”) es cometer petición de principio. El propio Putnam parece consciente de esa dificultad. Afirma explícitamente que su razonamiento depende de supuestos empíricos como “por ejemplo, el supuesto de que la mente no tiene otro acceso a las cosas y las propiedades externas que el que le proporcionan los sentidos” (Putnam 1981, p. 16). Ese supuesto desempeña un papel similar a P2 en su demostración. Por tanto, es razonable suponer que Putnam aceptaría lo siguiente: un sujeto que creyera que es un cerebro en una cubeta tendría una concepción sobre las relaciones entre nuestras representaciones y el mundo externo conforme a la cual no habría motivo alguno para creer tesis como P2 o como el supuesto empírico especificado en la cita anterior.

Para replicar a esa crítica –y poder apreciar mejor la fuerza y los límites del argumento putnamiano– conviene reflexionar sobre la relatividad del concepto de petición de principio. Se comete (o no) petición del principio *contra la posición epistémica de un sujeto*, el sujeto que sería el destinatario de la argumentación. Como sugiere nuestra descripción anterior (inspirada fundamentalmente en el análisis de Jackson 1984), Falvey y Owens están en lo cierto al sugerir que si proponemos (A) estamos incurriendo en petición de principio *contra alguien que crea que es un cerebro en una cubeta*. Sin embargo, ¿es ese sujeto el destinatario principal del argumento (A)?

Un hecho relevante –aunque no concluyente– es el siguiente: no existen (o al menos no tenemos el menor indicio de que existan) tales sujetos. No es concluyente porque podríamos imaginar que Putnam se proponía elaborar un ejercicio filosófico altamente especulativo e hipotético (más de lo usual en filosofía), conforme a una intención como ésta: “Supondré que un hipotético *x* cree que es un cerebro en una cubeta, y elaboraré un argumento que podría convencer incluso a *x* de que está en un error”. No obstante, podemos concebir un destinatario diferente del argumento (A), o del argumento original de Putnam: un sujeto normal y corriente, potencial lector de Putnam y otros epistemólogos, o –más específicamente– un miembro de la comunidad científico-filosófica. Ese sujeto no cree ser un cerebro en una cubeta. Probablemente cree incluso que no es un cerebro en una cubeta, y querría tener una idea sobre el estatus epistémico de dicha creencia. Se pregunta, por ejemplo, si dicha creencia constituye saber (como sería el caso si la creencia fuera la conclusión de un argumento correcto cuyas premisas conoce). Ofrecer el argumento (A) a ese sujeto (a cualquiera de nosotros) no es incurrir en petición de principio contra su posición epistémica.

## **Bibliografía**

- Falvey, Kevin y Joseph Owens (1994): “Externalism, Self-Knowledge, and Skepticism”, *The Philosophical Review* 103, pp. 107-137.
- Jackson, Frank (1984): “Petitio and the Purpose of Arguing”, *Pacific Philosophical Quarterly* 65, pp. 26-36.
- Pérez Otero, Manuel (2009): “Transmisión de la justificación epistémica”, en Daniel Quesada (coordinador), *Cuestiones de teoría del conocimiento*. Madrid: Tecnos, pp. 148-180.
- Pérez Otero, Manuel (2010): “Modest Skepticism and Question Begging Proper”. *Grazer Philosophische Studien* (próxima publicación).
- Putnam, Hilary (1981): “Brains in a Vat”. Capítulo 1 de su libro *Reason, Truth and History*, Cambridge: Cambridge University Pres.

# The open future: truth, bivalence and indeterminacy<sup>1</sup>

María Ponte Azcárate  
maria.deponte@ehu.es  
ILCLI, University of the Basque Country  
Villa Asunción, Jose Elosegui Etorb. 275  
20015. Donostia. Spain

Consider the following sentence<sup>2</sup>

(1) There will be a sea battle tomorrow.

Assume that nothing in the present state of the world determines whether there is going to be a sea battle tomorrow or not. There might, of course, be elements that make the event more or less probable, but none of them are fully conclusive. It seems then plausible to claim that (1), as uttered by me today,  $t_0$ , is neither true nor false. This is what has been called the *indeterminacy intuition*: since nothing in the present state of affairs determines whether the assertion is true or false, the assertion lacks truth-value.

Now, imagine we are at the next day,  $t_1$ , a sea battle is indeed taking place and we are witnessing it. In  $t_1$ , it seems I can claim,

(2) Sentence (1), as uttered by me yesterday, was true at the time of utterance,  $t_0$ .

This is what has been called the *determinacy intuition*: At the end of the next day,  $t_1$ , the sea battle has (or has not) happened and so my assertion of (1) was true (or false) at  $t_0$ . We are, generally, strongly tempted to accept both intuitions but, by doing so, we fall into the paradox of the future contingents.

MacFarlane (2003) has recently developed a relativist semantics that is able to incorporate both intuitions. Very briefly, he states that utterance-truth needs to be relativized. The idea is that, in order to deal with the paradox of future contingents, we need to take into account the so called “context of assessment”, the temporal perspective from which a sentence such as (1) is evaluated.

Accordingly, we would say that from the “context of utterance” (temporal perspective of when the assertion is made), it is right to say that (1) is neither true nor false (indeterminacy). And, from the “context of assessment” (temporal perspective of when it is evaluated), it is right to say that the utterance of (1) was either true or false (determinacy). MacFarlane does solve the paradox of future contingents, but the prize is very high. Accepting his solution would entail renouncing to some of the basic intuitions behind classical standard semantics.

In this paper, I will attempt to defend a solution to the open future problem according to which, roughly, future contingents are to be explained without having to relativize truth to the context of assessment and classical semantics is to be retained. In order to do so, I shall question some of the ideas implicit, and generally accepted, in the formulation of the problem of the open future.

Firstly, I will challenge the generally established idea that acceptance of bivalence (with regard to statements about the future) entails denial of indeterminacy or, in other words, acceptance of fatalism. The argument can be outlined as follows:

---

<sup>1</sup> This research has been partially supported by the Basque Government (IT323-10) and the Spanish Ministry of Science and Innovation (FFI2009-08574)

<sup>2</sup> This example is due to Aristotle. The formulation I’m presenting here is largely due to MacFarlane (2003)

1. Either it is true that there will be a sea battle tomorrow or it is true that there will not be a sea battle tomorrow
2. If it is true that there will be a sea battle tomorrow, then it is true NOW that there will be a sea battle tomorrow
3. If it is true NOW that there will be a sea battle tomorrow, or true NOW that there will not be, then how tomorrow is, is settled by how the present is
4. Since we are dealing with an arbitrary event at an arbitrary future time, how the future is in all respects is settled by how the present is.

Premise 3 is the one that needs to be rejected, for it states that if an event will come to be, then it is inevitable that it will come to be. In order to challenge this premise, we shall differentiate between two notions of truth: “be true” and “settled or determinate truth”<sup>3</sup>. The first is a tenseless property of propositions and is bivalent; the second entails a proposition being made true by events and is not bivalent. Once we accept this distinction, assuming that the future is open does not (necessarily) entail that the propositions expressed by utterances about the future lack truth value; but merely that they are *unsettled or ungrounded*.

The idea, roughly, is to say that every proposition is either true or false, but that for some propositions it is unsettled *which* truth-value they have. Determinately, it’s true that p or it’s true that not-p; but it’s neither determinate that p is true nor is it determinate that not-p is true. The world settles the matter that one or the other of a proposition and its negation is the case, but it fails to settle *which one* is the case.

This leads us to reformulate both the indeterminacy and the determinacy intuitions. With these new definitions at hand, I shall claim, it is possible to deal with the problem of the open future within classical semantics (à la Kaplan). Thus, future contingents are not to be taken as a reason to adopt any form of relativism. Granted, much more will be required to develop a fully convincing solution to the puzzle<sup>4</sup> and the view clearly raises some difficult metaphysical issues<sup>5</sup>, but it offers nonetheless an alternative and promising perspective on the problem.

## Bibliography

- Barnes E. and Cameron R. (2009) “The Open Future: Bivalence, Determinism and Ontology” *Philosophical Studies* 147, 2, 291-309
- MacFarlane J. (2003) “Future contingents and relative truth” *The Philosophical Quarterly*, 53, 321-336
- MacFarlane J. (2009) “Non-indexical contextualism” *Synthese*, 166, 231-250
- Perry, J. (2006) “How real are future events?” Stadler, F. and Stöltzer, M. (eds) *Time and History*, Ontos Verlag.
- Prior A (1957) *Time and Modality* Oxford: Clarendon Press
- Thomason R.H. (1970) “Indeterministic time and truth-value gaps” *Theoria* 36, 264-281

---

<sup>3</sup> Equivalent, to a certain extent, to the supervaluationist distinction between “truth on a history” and “truth simpliciter” or “determinately true”. See Perry (2006) for a similar distinction.

<sup>4</sup> Notably problems related with the acceptance of truth-value gaps and some counter-intuitive truth conditions that seem to be entailed by the proposal. (For instance, one drawback comes from allowing there to be true disjunctions or true existential generalizations that have no true disjunct or instance)

<sup>5</sup> Notably the so called problem of the “thin red line”. In order to avoid it, it seems we need to embrace either presentism or a shrinking view of time.

## Conventionalism and Metaontology

Laura Cecilia Porro  
lcp28@st-andrews.ac.uk  
Arché

Philosophical Research Centre for Logic, Language, Metaphysics and Epistemology  
The University of St Andrews  
17 - 19 College Street  
St Andrews  
Fife  
KY16 9AL  
United Kingdom

The purpose of this paper is to ask whether there is a connection between one's view about conventionalism and one's metaontological view. In order to explore this issue I will discuss two conventionalist approaches to metaphysics, Sider's [Sid01] and Sosa's [Sos03]. In the end I will argue that different conventionalist approaches lead to different metaontological positions.

In the first part of the paper I will briefly describe Sider's view of conceptual analysis as it is in his "Personal identity and the limits of conceptual analysis". I will show what it means for a metaphysical question to have a conventionalist answer and when, according to Sider, we can adopt a conventionalist approach towards a metaphysical question. I will follow Sider in using as a case study for this enquiry about conventionalism the debate about personal identity.

There are two types of metaphysical questions. Ontological questions (e.g. "are there mereological sums?") concern the existence of an entity. On the other hand non-ontological questions (e.g. "what are laws of nature?") do not concern the existence of an entity. I will present some remarks about the relevance of this classification. First of all, there are cases in which it is difficult to decide whether a question belongs to one category or another (e.g. "is there a thing such that it has as a criterion of persistence  $x$ ?"). Moreover, it is also controversial whether the distinction between ontological and non-ontological questions actually captures a relevant difference amongst metaphysical questions.

Sider claims that ontological questions cannot be answered by means of conceptual analysis. This means that ontological questions cannot have conventionalist answers and that there is a fact of the matter to determine the answer to an ontological question. On the other hand non-ontological questions can be answered through conceptual analysis and can thus have conventionalist answers. In the specific case of the debate about the criteria of personal identity, Sider claims that the question which criterion of personal identity is correct is a non-ontological question. This means that there is no fact of the matter to establish which criterion of personal identity is the correct one. On the other hand, there is a fact of the matter to settle the ontological question whether we should be 3-dimensionalist, 4-dimensionalist or nihilist.

The paper will try to answer the following questions:

1. If the debate about a non-ontological question is conventional, is this due to the fact that the debate about the underlying ontological framework is conventional as well? Is a conventionalist approach to answer one of the two types of question compatible with a non conventionalist approach to answer the other type of question?
2. How do vagueness, arbitrariness and indeterminacy arising in either case (i.e. non-ontological or ontological disputes) determine the answer we are going to give to other metaphysical disputes?
3. Do we use the same conceptual schemes in order to address both types of questions?

In order to answer these question I will discuss Sosa's paper [Sos03]. I will try to understand

whether Sosa's approach to ontological questions counts as conventionalism or not. Furthermore I will compare Sider's and Sosa's approach to ontological questions. The discussion of Sosa's work will help answering questions (1) and (3).

I will argue that Sider's and Sosa's views are incompatible. Sosa holds that we can adopt a conventionalist approach toward ontological questions, whereas Sider claims that ontological questions cannot have conventionalist answers.

In the second part of the paper I am going to show how different approaches to conventionalism can be related to different metaontological positions. We can draw four theoretical approaches to metaphysics, depending on their metaontological and metaphysical views. I will discuss the following four possible approaches:

- (a) metaontological dismissivism and use of conventionalism to address ontological disputes;
- (b) metaontological dismissivism and no use of conventionalism to address ontological disputes;
- (c) robust metaontological view and use of conventionalism to address ontological disputes;
- (d) robust metaontological view and no use of conventionalism to address ontological disputes.

In the last part of the paper I will briefly describe different types of vagueness. Then I will sketch how conventionalism is related to vagueness. I will state the claim that vagueness is the reason why we adopt conventionalism, but I will not argue for the claim and leave the issue open to further research.

#### References

[Sid01] Theodore Sider. Criteria of personal identity and the limits of conceptual analysis. *Philosophical perspectives*, 15:189–209, 2001.

[Sos03] Ernest Sosa. Ontological and conceptual relativity and the self. In Michael J. Loux and Dean W. Zimmerman, editors, *The Oxford Handbook of Metaphysics*. 2003.

# **Learning and genetic assimilation: A game theoretic interpretation of the Baldwin Effect**

Diego Rios  
Helsinki University (Finland)

Graciela Kuchle  
Witten Herdecke University (Germany)

## 1. ABSTRACT

The Baldwin effect is a process by which learnt traits become integrated into the genome through a non-Lamarckian mechanism. The Baldwin effect has been challenged on three grounds – its feasibility, plausibility and likelihood. The strategy that we follow in this paper is one of moderate vindication. First, we evaluate the arguments that have been purported to avert the mentioned criticisms. We present a game-theoretic perspective, capable of providing a more general perspective to evaluate the Baldwin effect.

## 2. THE BALDWIN EFFECT

An example may illustrate the mechanics of the process. Imagine that in a population of tits a new skill emerges – for instance, a trick to opening milk-bottles by pecking. How this new trait emerged is of no major significance. It could have been by sheer luck or by the clever discovery of an individual of the population. According to the usual account, the new skill spreads through the population by social learning. From that moment on, any genetic mutation facilitating the adoption of the trait will be selected. The upshot of the process is the genetic fixation of originally learnt skills.

## 3. WATKINS' UNFEASIBILITY ARGUMENT

In this section we deal with the unfeasibility argument raised by John Watkins (1999). To better understand Watkins' criticism we adopt Godfrey-Smith's characterization of the three stages in which the Baldwin effect takes place (Godfrey-Smith, 2003):

Stage 0: A new trait emerges in the population (a new technique for opening coconuts, for instance).

Stage 1: The new trait spreads through the population by learning.

Stage 2: New genetic mutations that facilitate the acquisition of the new trait are selected, contributing to the fixation of the originally adopted trait.

Watkins' objection is that the transitional stage 1 is either superfluous or committed to Lamarckism. According to Watkins' own words, the Baldwin effect assumes that "hitting on B (e.g. some adaptive behavior), as a successful solution to a recurrent problem, increases the probability that a gene  $\beta$  for B will get established. But why, and how? It would be Lamarckian to impute to the acquired characteristic B a propensity to engender  $\beta$ " (Watkins, 1999).

This is not the right way to uncover the structure of the mechanism governing Baldwinization. The issue is not that the adoption of behavior B increases the probability of assimilating gene  $\beta$ , but rather that  $\beta$  facilitates the appearance of adaptive behavior B both in the present and future generations.

We assume that adopters of the imprinting behavior (B) are selected with probability p and non-adopters of the behavior ( $\sim B$ ) with probability q. To avert Lamarckism, we assume that the initial probability of having the facilitating mutation,  $\beta$ , is the same for both adopters and non-adopters of B, namely,  $\Pr(\beta / B) = \Pr(\beta / \sim B) \equiv \lambda$ . Because the behavior B is adaptive, adopters are more likely to transmit their genes to the next generation ( $p > q$ ). Watkins acknowledges that mutations have better chances of being transmitted if they occur in adopters of B, and that it may in this sense be that the adoption of B fosters mutation  $\beta$ . Of course, this is uncontroversial. However, the crucial idea absent in Watkins's analysis is that  $\beta$  facilitates the emergence of B, and not that B raises the propensity of obtaining  $\beta$ . For the Baldwin effect to have a significant evolutionary relevance, it is only required that those individuals having mutation  $\beta$  have higher chances of displaying the adaptive behavior B than those individuals lacking the mutation  $\beta$ . The crucial conditional probability is the probability of displaying behavior B, given that mutation  $\beta$  is present i.e.  $\Pr(B / \beta)$ .

#### 4. THE POSITIVE FREQUENCY ACCOUNT OF THE BALDWIN EFFECT

Positive frequency dependence occurs whenever the advantage of having a trait increases with the proportion of individuals in the population who have a convergent trait. The idea of positive frequency dependence has much intuitive appeal. Nevertheless, we believe that it fails to pinpoint the crucial ingredients underlying Baldwinization. Note that positive frequency dependence is a form of increasing returns. Although several phenomena display this feature, there are many interactions characterized by constant and even decreasing returns to adoption. In the case of constant returns, the advantage of adopting a phenotype does not depend on the proportion of adopters and, in the case of decreasing returns, the advantage decreases with the frequency of adopters. The trick to opening coconuts provides an example of this last situation: the higher the proportion of individuals who have adopted the trick, the more difficult it will be to find coconuts to open. However, adopting the trick is adaptive even if it improves the fitness of the adopters at a decreasing rate. This example shows that the relevant criterion for adoption is not determined by the type of frequency dependence, but by the best-response nature of this behavior, given what other individuals are doing.

#### 8. CONCLUSION

We will conclude that, in stark contrast to mainstream approaches, the game-theoretic framework captures prototypic examples of Baldwinization that otherwise it would be difficult to taxonomize.

## La regla ECQ, el silogismo disyuntivo y las lógicas modales de Lewis

Gemma Robles

Dpto. de Hist. y Fil. de la CC, el Leng. y la Ed., Universidad de La Laguna. Facultad de Filosofía,  
Campus de Guajara, 38071, La Laguna, Tenerife.  
Correo-e: gemmarobles@gmail.com; URL: <http://webpages.ull.es/users/grobles>

La regla ECQ (“E contradictione quodlibet”) es:

$$\text{ECQ. } A, \neg A \vdash B$$

En lógicas donde las reglas  $E\wedge$  y  $I\wedge$ , i.e.,

$$E\wedge. A \wedge B \vdash A, B$$

$$I\wedge. A, B \vdash A \wedge B$$

son válidas, ECQ puede formularse de modo equivalente de la siguiente manera:

$$\text{ECQ. } A \wedge \neg A \vdash B$$

Como  $E\wedge$  y  $I\wedge$  son válidas en todas las lógicas de este trabajo, por ECQ entenderemos esta segunda versión de la regla.

Como es sabido (cf. [5]), una lógica  $S$  es *paraconsistente* si ECQ no es una regla de  $S$ . No menos conocido es el hecho de que las lógicas de Lewis no son paraconsistentes. Pero este hecho no es, desde luego, accidental. En efecto, Lewis sostiene que  $B$  es deducible de  $A \wedge \neg A$  (equivalentemente,  $A \wedge \neg A \rightsquigarrow B$  —  $\rightsquigarrow$  representa la implicación estricta) de acuerdo con el siguiente argumento ([4], p. 250) actualmente conocido como “prueba de Lewis” o “argumento de Lewis” (cf. [4], §16.1):

- |                      |                                  |
|----------------------|----------------------------------|
| 1. $A \wedge \neg A$ | Hyp.                             |
| 2. $A$               | $E\wedge$ , 1                    |
| 3. $\neg A$          | $E\wedge$ , 1                    |
| 4. $A \vee B$        | Introducción de la disyunción, 2 |
| 5. $B$               | Silogismo disyuntivo, 3, 4       |

El objetivo del presente trabajo no es examinar de nuevo este discutido argumento (cf., e.g., [2], [3]) sino definir una serie de lógicas paraconsistentes incluidas en la lógica  $S4$  de Lewis. En este sentido, obsérvese que en el “argumento de Lewis” el silogismo disyuntivo (s.d) se usa como una *regla de inferencia*, i.e., en la forma:

“De  $\neg A$  y  $A \vee B$  infiérase  $B$ ”.

En otras palabras, se usa como una regla susceptible de ser aplicada a cualesquiera premisas. Considérese, sin embargo, s.d en la siguiente forma:

“Si  $\neg A$  y  $A \vee B$  son derivables, entonces  $B$  es derivable”.

Entendida de este modo, s.d es una *regla de prueba*: sólo es aplicable a teoremas. Y si s.d es una regla de prueba, entonces, claramente, la prueba de ECQ de Lewis no se sostiene.

Dado que en todas las lógicas de este trabajo s.d es entendida como una regla de prueba, ECQ no es válida en ninguna de ellas. Las teorías construidas sobre estas lógicas, por lo tanto, no están necesariamente cerradas por s.d, y consiguientemente, en caso de inconsistencia, no tienen por qué devenir triviales forzosamente. No obstante, las teorías primas consistentes siempre están cerradas por el silogismo disyuntivo, por supuesto. La serie de lógicas a la que me he referido más arriba comenzará con la lógica básica  $B$  de Routley y Meyer complementada con el axioma característico de  $S4$

$$B \rightarrow (A \rightarrow A)$$

y las reglas *reductio*

“Si  $A \rightarrow \neg A$  es derivable, entonces  $\neg A$  es derivable”

y s.d. Esta lógica será denominada  $B_{K2\gamma}$  donde B es la lógica B de Routley y Meyer, K2 es la siguiente restricción (a teoremas implicativos) de la regla K:

“Si  $A \rightarrow B$  es derivable, entonces  $C \rightarrow (A \rightarrow B)$  es derivable”

y  $\gamma$  es la denominación con la que Ackermann se refería al silogismo disyuntivo (cf. [1]).

Posteriormente,  $B_{K2\gamma}$  será extendida con algunas tesis conocidas hasta una restricción de S4 que puede describirse de diversas (y equivalentes) maneras desde un punto de vista intuitivo. Por ejemplo:

a) S4 sin las reglas *reductio* como, e.g.,

“Si  $A \rightarrow B$  es derivable, entonces  $(A \rightarrow \neg B) \rightarrow \neg A$  es derivable”

o

“Si  $A \rightarrow B$  es derivable, entonces  $(\neg A \rightarrow B) \rightarrow B$  es derivable”.

b) S4 sin la regla K

K. “Si  $A$  es derivable, entonces  $B \rightarrow A$  es derivable”

o, equivalentemente, sin la regla *aserción*:

“Si  $A$  es derivable, entonces  $(A \rightarrow B) \rightarrow B$  es derivable”.

c) S4 sin el axioma ECQ

$$(A \wedge \neg A) \rightarrow B$$

Todas las lógicas tratadas en este trabajo contienen la lógica clásica en el sentido de que todas las tautologías en  $\wedge$ ,  $\vee$  y  $\neg$  son derivables. Es más, como s.d es una regla primitiva de cada una de ellas, disponemos del modus ponens para  $\supset$ .

Daré por supuesto que se conoce la semántica de Routley y Meyer para las lógicas de la relevancia.

#### *Agradecimientos*

- Trabajo financiado por los proyectos de investigación FFI2008-05859/FISO y FFI2008-01205/FISO del Ministerio de Ciencia e Innovación de España.
- G. Robles está respaldada económicamente por el Programa Juan de la Cierva del Ministerio español de Ciencia e Innovación.

#### REFERENCIAS

- [1] Ackermann, W., “Begründung einer strengen Implikation”, *Journal of Symbolic Logic*, 21/2, 113-128, 1956.
- [2] Anderson, A. R., Belnap, N. D. Jr., *Entailment. The Logic of Relevance and Necessity*, vol I. (Princeton University Press), 1975.
- [3] Anderson, A. R., Belnap, N. D. Jr., Dunn, J. M., *Entailment. The Logic of Relevance and Necessity*, vol II. (Princeton University Press), 1992.
- [4] Lewis, C. I., Langford, H., *Symbolic Logic*, Dover, New York, 1959, 1932.
- [5] Priest, G., Tanaka, K., “Paraconsistent Logic” (in E. N. Zalta ed.), *The Stanford Encyclopedia of Philosophy. Winter 2004 Edition*.  
URL: <http://plato.stanford.edu/archives/win2004/entries/logic-paraconsistent/>.

## **Género y Nuevas Tecnologías. Los estereotipos de género también navegan por el ciberespacio**

Amparo Romero Sánchez  
Universidad de La Laguna

Actualmente es imposible negar la importancia de la ciencia y la tecnología en las vidas de las personas. La tecnociencia se ha convertido en una fuente de poder que domina todos los ámbitos de la sociedad: la educación, el mundo laboral e, incluso, la política. Hoy más que nunca el índice de progreso de una sociedad se basa de forma preeminente en sus logros científicos y tecnológicos. Pero la ciencia y la tecnología también son fuente de peligros y de nuevas desigualdades.

Las nuevas tecnologías de la información y la comunicación (TIC's) suponen para muchas personas una ruptura con el carácter de los desarrollos tecnológicos anteriores. Las TIC's han acarreado, junto con un buen número de beneficios una nueva fuente de desigualdad que tiene lugar cuando las personas quedan al margen del desarrollo tecnocientífico y, por lo tanto, quedan excluidas de los ámbitos de decisión y de poder.

Todos estos análisis responden a la urgente necesidad de establecer el impacto que tienen las nuevas tecnologías sobre la sociedad en su conjunto y, en particular, sobre las vidas de las personas, estableciendo las necesarias contextualizaciones mediadas por diversas categorías (género, país o región de origen, edad, etc.) a la hora de realizar los análisis.

Las actuales reflexiones, que suponen un profundo cuestionamiento de la relación entre tecnociencia y sociedad, hallan sus antecedentes en la década de los sesenta con el comienzo de la revisión crítica de la imagen tradicional de la ciencia y tecnología gracias a la confluencia de varios factores como, por ejemplo, las protestas de determinados movimientos sociales, que hicieron posible el cambio de percepción hacia dichos ámbitos. Hasta ese momento, la visión que predominaba afirmaba la autonomía, la neutralidad y el progreso lineal de dichos campos del conocimiento.

La revisión crítica de la ciencia fue abordada desde diferentes perspectivas, naturalistas, sociológicas, económicas, etc.. Actualmente, está protagonizada en el ámbito académico por lo que se ha denominado estudios CTS, es decir, estudios de ciencia, tecnología y sociedad. Los estudios CTS no constituyen un campo homogéneo de conocimiento, pues está formado por diversas disciplinas (sociología, filosofía, economía, etc), aunque comparten y parten en sus análisis de una premisa básica: la ciencia y la tecnología no están aisladas de las sociedades en las que surgen, sino que, al contrario, su dinámica responde, además de a características internas relativas a dichos ámbitos, a procesos sociales exteriores a ellos.

Los estudios CTS han desembocado actualmente en las propuestas constructivistas, que enfatizan la dimensión social de la ciencia y la tecnología. Estas propuestas rechazan la percepción de la ciencia como una actividad intelectual pura y la de la tecnología como ciencia aplicada. En el caso concreto de la tecnología, defienden la idea de que la interacción con el ámbito social tiene lugar a través de relaciones recíprocas en las que la tecnología modifica las relaciones sociales y éstas, a su vez, influyen en el desarrollo tecnológico.

Esta perspectiva constructivista ha sido bien acogida por el feminismo, movimiento que, además, la ha completado, aportando el género como una categoría de análisis relevante para reflexionar acerca de las relaciones entre tecnología y sociedad. Esta alianza entre constructivismo y feminismo es la postura adoptada en el presente

trabajo, en el que se pretende demostrar que la tecnología en general y las nuevas tecnologías en concreto, lejos de ser una actividad neutral, es reflejo de los valores de la sociedad en la que se desarrolla.

En el tema del impacto tecnológico en la vida de las mujeres, el feminismo ha estado dividido tradicionalmente en dos posturas, a saber ¿utopía o distopía? ¿La tecnología como instrumento de dominación sobre las mujeres o como elemento propiciatorio de la emancipación de las mujeres?

Por un lado hallamos, pues, las corrientes feministas que abogan por el desarrollo tecnológico al ver en él un camino de liberación para las mujeres, y las corrientes feministas que, al contrario, perciben la tecnología como un proyecto inherentemente patriarcal encaminado a la dominación y explotación de las mujeres.

En el feminismo de las décadas de los 70 y 80 la tendencia general de las feministas con respecto a la tecnología fue la del pesimismo. Como ejemplo, tenemos al feminismo radical de la época, movimiento que consideró que el proyecto tecnológico va encaminado a afianzar la dominación de las mujeres y de la naturaleza.

En la actualidad, la tendencia general con respecto a las nuevas tecnologías es el optimismo. La corriente feminista de pensamiento que ejemplifica de manera magistral esta postura esperanzadora hacia las nuevas tecnologías es, sin duda, el ciberfeminismo que se caracteriza por el optimismo utópico con respecto a las nuevas tecnologías, a las que las ciberfeministas consideran esencialmente liberadoras para las mujeres por su carácter flexible y abierto que, desde su punto de vista, permite la creación libre de la identidad.

El propósito es la defensa de una postura intermedia que sea consciente de las ventajas que ofrecen las nuevas tecnologías pero que, al mismo tiempo, las acepte de una manera crítica, utilizando la categoría del género para dilucidar los sesgos y valores patriarcales que están presentes en la tecnología.

La perspectiva protagonista se denomina “tecnofeminismo” y constituye una alianza entre constructivismo y feminismo. Se trata, en definitiva, de un constructivismo con perspectiva de género que posibilita que la relación entre tecnología y género sea percibida como “mutuamente conformadora”<sup>1</sup>.

Aplicando la perspectiva mencionada, se ha analizado la tendencia reciente de la creación de páginas webs dedicadas exclusivamente a “juegos para chicas” o “juegos para niñas”. En este tipo de páginas el contenido, desde mi punto de vista, no responde a demandas y necesidades reales de las “chicas”, sino a estereotipos de género tradicionales que hacen acto de presencia en el proceso de diseño del contenido de los mencionados juegos. En este sentido, podemos hallar páginas en Internet en las que puede observarse que los “juegos para niñas” o los “juegos para chicas” constituyen un espacio feminizado dentro del masculinizado mundo tecnológico del ciberespacio.

El aspecto relevante de dicho análisis es, en definitiva, que la creación de espacios feminizados dentro del mundo tecnológico reafirma el carácter masculino de las nuevas tecnologías, puesto que nos dice que las mujeres deben ver un toque “femenino” en la tecnología para interesarse por ella. El factor clave es el hecho de que el diseño y el contenido cambian cuando se trata de producir tecnología destinada al consumo femenino reforzando, con esto, la separación de esferas de hombres y mujeres.

---

<sup>1</sup> J. Wajcman, *Tecnofeminismo*, Ed. Cátedra, Madrid, 2004, p. 161.

## **Anomaly in Novel Metaphor and Experimental Tests**

Esther Romero

Facultad de Psicología, Departamento Filosofía I

Campus Cartuja, 18071 Granada

[eromero@ugr.es](mailto:eromero@ugr.es)

and

Belén Soria

Facultad de Filosofía y Letras, Departamento de Filologías Inglesa y Alemana

Campus Cartuja, 18071 Granada

[bsoria@ugr.es](mailto:bsoria@ugr.es)

Our aim is to argue for contextual abnormality as a necessary condition of the metaphor identification criteria taking into account neurological experimental evidence on metaphor processing.

That metaphor identification is a part of metaphor processing is a proposal, nevertheless, that many cognitive metaphor theorists have rejected. Their most persuasive attack, based on empirical evidence from reaction time experiments, is their argument against any version of anomaly. Although this argument takes different forms depending on the particular theoretical proposal under consideration, it always incorporates the premise according to which there is identity of reaction times in interpreting literal and metaphorical uses of language. In addition, it always incorporates a rejection of two standard pragmatic proposals on metaphor as implicature. In particular, with this argument, they express disapproval of (a) that an anomaly (understood as categorial falsity) is present in the identification of metaphorical utterances and (b) that the interpretation of metaphorical utterances proceed in two propositional stages, first they are interpreted literally and then metaphorically.

For cognitive metaphor theorists (Ortony et al. 1978, Gerrig 1989, Gibbs and Gerrig 1989, Keysar and Glucksberg 1992, Gibbs 1994), the argument for the rejection of these two theses is as follows. They assume that the categorial falsity criterion entails the two propositional stages in metaphorical interpretation and that the two propositional stages in metaphorical interpretation, in its turn, entails that there should be an additional time in metaphorical interpretation. As this last consequence is false because certain psychological reaction time experiments show that the interpretation of literal and metaphorical utterances exhibit equivalent processing time, the two propositional stages in metaphorical interpretation must be rejected and this involves the rejection of the categorial falsity criterion as well. Many cognitive metaphor theorists go further and also claim that any identification criterion of metaphor should be avoided. This conclusion depends on an additional assumption that they believe justified by the equivalent processing time for both literal and metaphorical interpretation. In particular, it depends on the assumption that the same procedures are used in the literal and the metaphorical interpretations. If this is so, it does not make sense to look for an identification criterion that triggers a metaphorical interpretation.

Although we agree with the rejections of both (a) and (b) (our reasons to reject both proposals will be considered in section 1), we consider the first part of the previous argument unsound because the processing time identity is questionable on a closer inspection of psychological experiments. Reaction time experiments do not always support reaction times identity for metaphorical and literal interpretations (Giora 1997).

Thus, the results of reaction time experiments, being inconsistent, are not evidence for anything: neither for the rejection of (a), anomaly as categorial falsity, nor for the rejection of any other proposal on metaphorical identification.

Nevertheless, if there were reasons to admit just the experiments supporting reaction times identity, we could reject (a), the anomaly as categorial falsity, but not any conception of anomaly that does not involve interpretation in two propositional stages (any subpropositional conception of the anomaly, Kittay's incongruity and our contextual abnormality, does not necessarily involve interpretation in two propositional stages) since the part of their argument against any type of anomaly counts with several unargued assumptions that are unjustified. In particular, we consider unacceptable the assumption that equivalent processing times entail equivalent effort. It is possible to take the same amount of time to lift a 5-pound and a 20-pound weight but the latter requires more effort. We also find unacceptable that for processing time identity we must conclude that metaphorical and literal interpretations demand exactly the same cognitive processes. It is possible to perform two different tasks at the same time even when these tasks are of a very different kind.

The situation created by the inconsistent results of the reaction time experiments together with our rejection of some unargued assumptions permit us to rescue virtually any theoretical conception about anomaly.

Although it is not easy in pragmatics to find crucial experimental evidence that clearly confirms one claim and disconfirms another (Sperber and Novek 2004), at least in the topic we are considering not everything is lost. Contextual abnormality could be preserved taking into account more recent neurological empirical studies on metaphor processing designed specifically to test hypotheses related with anomaly. According to them, there is anomaly in novel metaphor, marked in the brain with a larger amount of bilateral frontal and temporal cortex, and this anomaly appears in metaphor without two propositional stages in its interpretation (Tatter et al. 2002, Ahrens et al. 2007). Our contextual abnormality is a version of anomaly that, as we will show, is compatible with these more recent and more specific experimental findings.

We structure this talk in five sections. In section 1, "From semantic anomaly to contextual abnormality", we expound the metaphorical identification criterion, the semantic anomaly, of the standard pragmatic view of metaphor as implicature and explain how some modifications to this criterion can be introduced to avoid some of the problems posed to the semantic anomaly. In section 2, "Negation of anomaly in psycholinguistics", we consider in detail how the result of psychological experiments seems to encourage the dissolution of any identification criterion and analyze the interpretation of the results of the psycholinguistic experiments on metaphor and the experimental methods used in them. In section 3, "Rescuing anomaly", we rescue the proposal of anomaly by showing the unreliability of the experiments. There should be an assessment on how to handle the fact that the results of recent reaction time experiments oppose the previous ones. In addition, we rescue some versions of the anomaly unmasking the fallacy that allows them to get to the conclusion that, in the interpretation of the literal and the metaphorical use of language, the same procedures are used. In section 4, "ERP and functional MRI experiments support contextual abnormality", we explore more empirical evidence to get out of the situation created by the inconsistent results derived from reaction time experiments and to prove that there is anomaly in metaphor. Thus, we conclude, in section 5, that our contextual abnormality not only solves the problems that other anomaly proposals have but also is compatible with all experimental results.

## Truthmaker Theory Without Truthmakers.

Pablo Rychter  
University of Girona & Somerville College (Oxford)  
pablo.rychter@udg.edu

The idea that there may be truthmaking without truthmakers is familiar. Several authors have argued that although every true proposition is made true by reality, there need not be particular entities (like facts, states of affairs, or tropes) that make these propositions true. Here I would like to go further and argue that we can also have *truthmaker theory* (or an important part of it) without truthmakers. I will argue that to an important extent, talk about truthmakers within the theory can be understood as metaphoric and as conveying valuable information about something other than the supposed truthmakers (about *truthmaking*). This *figuralist approach*, as I will call it, is intended to be conciliatory between friends and foes of truthmaker theory. On the one hand, the approach acknowledges the importance of truthmaker talk, and shows why critics of truthmaker theory should take note of the theory even if they are right and there are no truthmakers. On the other hand, the figuralist approach acknowledges the force of the arguments for the view that truthmaking may obtain without truthmaker entities, and that belief in these entities is unmotivated. But the figuralist approach may be resisted by those who adopt a ‘superficialist’ meta-ontology according to which the existence of states of affairs follows trivially from facts about their alleged components. I shall argue that the reasons that seem to support this superficialist meta-ontology can be accommodated within the figuralist approach.

### 1. Truthmaking without truthmakers.

Truthmaker theory is based on the following fundamental principle:

(TM) Necessarily, for any proposition  $p$ , if  $p$  is true then there is some entity  $e$  in virtue of which  $p$  is true.

But what does it take for  $p$  to be true in virtue of  $e$ ? Most truthmaker theorists agree that at least part of what it takes for  $p$  to be true in virtue of  $e$  is for  $p$  to be *necessitated* by  $e$ , in the following sense: necessarily if  $e$  exists,  $p$  is true. This is captured by the thesis known as *truthmaker necessitarianism*, which has some interesting ontological consequences. Although it allows ordinary concrete individuals to be truthmakers for some truths (Obama, for instance, is a truthmaker of <Obama exists>), it implies that the truthmakers of most contingent truths are something other than ordinary, concrete individuals. The proposition <Obama is smart>, for instance, is not made true by Obama: since Obama could fail to be smart, he does not necessitate the truth of the proposition. So in order to ground the truth of <Obama is smart>, most truthmaker theorists appeal either to the state of affairs *Obama's being smart* or to the trope *Obama's smartness*. Given some extra assumptions about the nature of these entities, they necessitate the truth of <Obama is smart>, as required by (TM). Thus, truthmaker theorists typically admit in their ontology either states of affairs or tropes.

But this is one point where some critics of truthmaker theory raise their eyebrows: unless we have independent reasons for believing in tropes or states of affairs, the adoption of (TM) comes at a high ontological price. And this price, the critics think, is not worth paying, since the benefits of adopting (TM) can be achieved by a weaker principle that is ontologically innocuous. The basic motivation behind (TM) is that it captures the intuition that *truth is grounded in reality* or, in other words, that truths are true *because*, or *in virtue of* how the world is. But this motivation can be captured equally well, or so the critics argue, by something like the following principle:

(TM\*) True propositions are true in virtue of what things exist and how they are.

Unlike (TM), (TM\*) does not require the existence of tropes or states of affairs. The proposition that Obama is smart, for instance, is true because Obama exists and is smart. In sum, the intuition that truth is grounded in reality recommends (TM\*) and not (TM). Those who commit themselves to the truth of (TM\*) but not to the truth of (TM) can be said to believe in *truthmaking without truthmakers* (I call this view *TWT* for short).

I am sympathetic to *TWT* and I share the impression that truthmaker theorists are too ready to incur in new ontological commitments. But this impression rests on the assumption of a particular view of the nature of existence claims, one that the truthmaker theorist need not share. The view in question, which I will call '*ontological realism*', is the view that the existence questions that philosophers are concerned with are very often substantive and answerable only after many empirical and quasi-empirical considerations. But the truthmaker theorist may have different meta-ontological beliefs. He may be some kind of 'superficialist' about existence questions and argue as follows: 'the state of affairs *Obama's being smart* is a free lunch that you have if Obama exists and is smart. It is a different entity from Obama and from the property of being smart, but in an important sense it is not a *further* entity, nothing *over and above* Obama and the properties he has. So there is no point in resisting (TM) while adopting *TWT*: by claiming that Obama is smart, as *TWT* theorists do in order to explain the truth of <Obama is smart>, they are thereby committed to the corresponding state of affairs.' So the truthmaker theorist's readiness to embrace states of affairs can perhaps be explained by his assumption of this kind of superficialist meta-ontology.

In the first part of this talk, I will sketch a figuralist approach to truthmaker theory and argue for its initial plausibility. In the second part, I defend the approach from the threat of a superficialist meta-ontology.

## **Cuatro Tipos de Emergentismo**

Marcelo H. Sabatés

Philosophy Department

Kansas State University

El emergentismo acerca de lo mental, y mas generalmente, acerca de niveles ontológicos no básicos, se ha entendido usualmente como un intento de mantener una ontología naturalista evitando los excesos del reductivismo. Autores contemporáneos como Brian McLaughlin, Jaegwon Kim y Achen Stephan, entre otros, rastrean el emergentismo hasta el *System of Logic* de Mill e incluyen en él a autores como Morgan, Alexander y Broad. A ellos se suman filósofos mas recientes como Bunge, Searle y O'Connor, entre otros. Este trabajo se propone, primero, clarificar una serie de tesis que podrían atribuírsele al emergentismo y delinear cuáles de ellas marcan los límites de cualquier emergentismo. A continuación se reconstruyen (con cierta base histórica e interpretativa reciente) cuatro tipos de emergentismo posible. Finalmente se argumenta que sólo una de las posiciones cumple con la caracterización usual del emergentismo, pero dicha posición resulta una versión claramente deflacionista.



## Quantification does not involve identity

Francisco Salto  
Universidad de León  
Facultad de Filosofía y Letras  
Campus Vegazana 24071  
[francisco.salto@unileon.es](mailto:francisco.salto@unileon.es)

"There might be a problem about how to define identity to someone sufficiently lacking in conceptual resources, ... but since such unfortunates are rare, even among philosophers, we needn't worry much if their condition is incurable" David Lewis

This communication presents an argument from such an unfortunate, showing how nonstandard identity notions are theoretically grounded. Such alternative notions may happen to be also practically useful or even truthful, but here it is only their theoretical coherence under discussion.

1. Cardinal and predicative notions of identity are distinguished.
2. Reasons for undefinability of identity are reviewed in first and second order formal languages
3. Structures without numerical identity are proved to exist in general, as based on sets and other collections
4. Two relevant examples with philosophical significance are considered: finite unordered collections and purely relational domains
5. Oxonian counter arguments are

1. Uniqueness and sameness are two distinct relations. To clarify their distinction, we talk about cardinal and predicative notions of identity. Two readings in the article from our natural language are distinguished: its *cardinal* and its *presuppositional* reading. An occurrence of an article has a cardinal reading when it is marked to express some number (in particular, the number 1 in the case of cardinal uses of the determinate article). In other case, its reading is presuppositional. For example in (i) "the" has a presuppositional reading, while in (ii) it is cardinal:

- (i) the innocent should throw stones
- (ii) the last one close the door

A general outline of the criteria backing this distinction are given. Let's distinguish notationally between the cardinal and the presuppositional use by writing between squarebrackets "[...]" any article to be read cardinally. In this way we shall distinguish between:

\_\_\_ is the same as \_\_\_  
and

\_\_\_ is [the] same as \_\_\_

Let's call *uniqueness* the notion expressed by the predicate "being exactly one" and call *sameness* the notion expressed by the predicate "being itself". We may state now the STANDARD THESIS of identity, namely,

$x$  is identical to  $y$  iff  $x$  is [the] same as  $y$

(i) uniqueness and sameness are conceptual ingredients of the notion of identity, in the sense of expressing necessary and sufficient conditions for identity. (ii) Both notions are not synonymous nor logically equivalent, nor reducible in general one to the other. (iii) we hypothesize that both exhaust the meaning of identity, in strict uses of it (not in loose ones, which are by the way ubiquitous).

We claim that not all domains are standardly identical, nor all structures over them. That is, domains in which  $x$  is not [the] same as  $x$ . Hence the standard thesis is not trivial or truistic. As a consequence, there are genuine problems about identity and not only problems stated in terms of identity.

2. A previous concern is whether any rational or discursive activity whatsoever involve or presuppose identity, in which case quantification would. Three reflections answer this doubt:

- alliteration and its recognition does not presuppose numerical identity
- repetition of instances require only weaker relations than identity
- consider the radical case of Paul, a reasoner whose logical consequence operator is not reflexive. He is still able to succeed with loose identificatory tasks

3. A usual way of understanding identity is, informally, the relation that each thing bears to itself and to nothing else (Williamson, 2006; Hawthorne, 2003). It corresponds to the diagonal function over the domain, which is often understood in terms of naive set theoretical inclusion.

We shall argue both informally and formally that the diagonal of the domain does not identify up to uniqueness, since “being identical to itself” only ensures sameness. The case of undistinguishable Urelemente is considered.

Moreover, the arguments for the unaxiomatizability and undefinability of identity in first order are reviewed and the second order case discussed.

4. A more general way of understanding identity is as invariance under automorphisms (Béziau, 2004, Krause and Coelho 2002). The notion is defined and explained, including the distinction between rigid and non rigid structures (in which sameness is not the only automorphism).

We argue again that this notion does not involve identity showing the existence of models that are not standardly identical (according to the thesis above). In particular, we consider structures based on collections or domains that may not be sets. In general, [the] only value identical to a given value may not exist, even if the only value identical to itself is such value.

5. Examples and applications. Two philosophically relevant scenarios are exhibited, where the availability of nonstandard identity criteria is not superfluous. The first example is that of unordered finite collections. Sorites paradoxes for one part and finitistic simulations of infinitude on the other are two contexts where such unorthodox collections naturally appear (even if both are disputable). The second example is that of a purely relational domain, that is, a domain in which relata are defined as relations.

6. Each for different reasons, both Evans and Dummett seemingly disagree with the thesis here defended. They think that predication (Evans) or quantification (Dummett) involve identity, in the sense of refuting inscrutability theses (Evans) or relativity theses (Dummett). It is shown that weaker (non-uniqueness implying relations) are suited to the function both attribute to identity. We illustrate how Paul (the forementioned case) can manage to learn or –in the proper sense- consistently enlarge given predications and quantifications even without standard identity.

## REFERENCES

- M. Dummett, “Does quantification involve identity?”, in H.A. Lewis, *P. Geach: Philosophical Encounters*, London, Deventer, 1991. Republ. In *The Seas of Language*, Oxford, Clarendon, 1996
- M. Dummett, “Wang's paradox” **Synthese** 30 (1975) 301-324
- D. Krause, *The Mathematics of Non-individuality*, IEA, Sao Paolo, 2003
- D. Krause and A. Coelho, “Identity, Indiscernibility and Philosophical claims”, Ms. 2002
- J.Y. Béziau, “Identity, Structure and Logic”, **Bulletin of the Section of Logic** 25 (1996), 89-94
- G. Evans, “Predication and Identity”, **Journal of Philosophy** 72 (1975) 343-363

## Sobre la función de los pronombres anafóricos en el lenguaje natural

Aránzazu San Ginés / aransangines@gmail.com

Universidad de Granada / IHPST (Paris 1)

La anáfora es uno de los fenómenos lingüísticos que más amplia y persistentemente ha llamado la atención de lógicos, filósofos del lenguaje y lingüistas. El objetivo del presente trabajo es el de proponer una nueva perspectiva desde la cual tratar una de las muchas incógnitas que plantea este fenómeno. Nos interesan en particular los pronombres anafóricos, esto es, un tipo de pronombre cuya interpretación o aportación semántica al discurso viene enteramente determinada por otras expresiones del mismo. La pregunta a la que nos disponemos a dar respuesta es la siguiente: ¿Cuál es el papel que juegan estos pronombres en el contexto del lenguaje natural?

Observemos por un momento los siguientes ejemplos en los que aparecen pronombres anafóricos:

- (1) *No **prudent man** will drive when **he** is drunk.*
- (2) ***Few MPs** came to the party, but **they** had a marvellous time.*
- (3) *A man who put **his paycheck** in the bank is wiser than a man who put **it** in the BUECU.*
- (4) ***A man** loved **a woman**. **He** kissed **her**.*

Estos ejemplos presentan una característica que les hace especialmente relevantes: cada pronombre parece recuperar de manera diferente su antecedente. En particular, cada uno de los pronombres que aparecen en los tres primeros enunciados parece recuperar una ‘dimensión’ específica de su antecedente: en (1) *he* recupera una variable (el pronombre parece comportarse como una variable ligada), en (2) *they* recupera una referencia (el pronombre parece comportarse como un término singular) y, finalmente, en (3) el pronombre *it* recupera una expresión lingüística (comportándose pues como un ‘pronoun of laziness’).

Al parecer, si existiera un tratamiento uniforme de los pronombres anafóricos, este no podría ser ninguno de los anteriores. Nuestro objetivo durante la ponencia será mostrar que un tratamiento alternativo y uniforme es posible (e incluso natural). Para ello partiremos fundamentalmente de las ideas de dos autores: P.T.Geach y F. Corblin.

Una primera aproximación esperanzadora a la cuestión la encontraremos de la mano de Corblin (2006). Este autor se refiere aquí explícitamente a *una teoría uniforme de los pronombres entendida como ‘eco de menciones’*. Lo que defiende es que la función, uniforme, que realizan los pronombres en el discurso es la de repetir dentro de sus contextos específicos, la aportación que hacen a su vez sus antecedentes en sus respectivos contextos. Corblin nos da pues una primera pista. Pero ¿cómo traducir sus ideas desde el punto de vista formal? Será entonces cuando volvamos la vista a Geach (1962). Según este último, la principal función de los pronombres anafóricos es la de colaborar en la articulación de los discursos a través de la posición que estos ocupan en los distintos predicados (lo que permite, por ejemplo, formar un predicado unario a partir de uno binario: *to stand in the relation F to himself*).

Mi propuesta, construida sobre la base de estas ideas, es la siguiente: El pronombre recupera del antecedente una parte de la ‘estructura’ en la que éste se encuentra enclavado, modificando o añadiendo nuevos elementos a los ya existentes. Es decir, el

pronombre recupera la *mención* o uso que se dio al antecedente, recuperando el 'contexto estructural' en el que este último se encuentra incluido. Pero ¿qué es lo que se identificará formalmente como 'estructura del discurso'? En cuanto a esto, propondremos una representación diagramática del discurso, en virtud de la cual identificaremos el hecho de recuperar una *mención* con el de recuperar una parte de la representación diagramática. Pero ¿por qué decantarnos por el uso de diagramas?

En primer lugar, la utilización de diagramas hace posible el acceso y representación compacta de toda una globalidad de informaciones. El trabajo con diagramas hace la manipulación de datos mucho más ágil e informativa. De ahí que defendamos como ventajosa toda propuesta que suponga una representación en parte diagramática del lenguaje. También atribuiremos una cierta naturalidad a la propuesta diagramática. El papel fundamental que parece jugar la habilidad de *ver* en los seres humanos, ese *talento innato* del que nos habla Hoffman (2000), ha sido cuidadosamente desarrollado y refinado a lo largo de la historia evolutiva del hombre. Parece pues natural, e incluso esperable, que el hombre se haya servido de este talento en procesos que le son especialmente propios, como lo son la construcción racional y el lenguaje. Los seres humanos razonan, se comunican en un mundo de imágenes. ¿Por qué deberíamos pues asumir que la visión no juega ningún papel substancial en ninguno de estos dos procesos?

A continuación, procederemos en la ponencia a describir algunas de las reglas de traducción (del lenguaje natural al de diagramas) así como la semántica del nuevo sistema a través de algunos ejemplos.

En lo concerniente a la semántica nos comprometeremos con tres principios:

1. El objetivo de un ser humano que emplea el lenguaje es el de comunicar rápido y de manera efectiva antes que el de comunicar con toda precisión. La posibilidad misma de error durante el discurso no constituye un hecho extraordinario e indeseable sino un aspecto de pleno derecho del proceso comunicativo.

2. El modelo no está fijado de antemano. Los enunciados aportan sucesivas restricciones, dando lugar a un modelo que se va construyendo paso a paso (los enunciados dan pistas).

3. La interpretación del discurso depende directamente de su representación diagramática. Más aún, este aporte de la estructura a la significación es dinámico. Es decir, la representación e interpretación de cada nuevo enunciado del discurso depende del diagrama que lo precede.

Por último, para concluir la ponencia recuperaremos los ejemplos del principio (1), (2), (3) y (4) y mostraremos que la funcionalidad que hemos atribuido de manera uniforme a los pronombres anafóricos se aplica bien a cada uno de ellos.

Corblin, F. (2006), "Pronouns and mentions". In *Comparing Anaphors. Between Sentences, texts and Languages*, I. Korzen, and L. Lundquist (eds.), Copenhagen Studies of Language 34, pp.27-43.

Geach, P.T. (1962), *Reference and Generality. An Examination of Some Medieval and Modern Theories*. Cornell University Press.

Hoffman, D.D. (2000), *Visual intelligence: how we create what we see*, W.W. Norton & Company, New York – London.

## Ontological versus Ideological Hierarchies

Gonçalo Santos

(Logos, University of Barcelona and LanCog, University of Lisbon)

goncalo.b.santos@gmail.com

LOGOS - Dep. Lògica, Història i Filosofia de la Ciència

Universitat de Barcelona, Facultat de Filosofia, Despatx 4023

Carrer Montalegre 6-8, Barcelona 08001

Several authors (David Lewis (1991), Vann McGee (2000) and Timothy Williamson (2003)) have argued that when putting forward his thesis, the generality relativist automatically defeats himself. The main goal of this paper is to see whether the generality absolutist is in a much better position. In particular, we investigate whether adopting Rayo's (2006) proposed extension of a Boolos-style treatment of second-order quantification to third and higher-order quantification, can be employed to deal with a problem for generality absolutism first noticed by Williamson (2003).

By the standard notion of logical consequence we say that  $\alpha$  is a logical consequence of  $\Gamma$  if and only if  $\alpha$  is true in every interpretation that makes every sentence in  $\Gamma$  true. Notice that by the notion of logical consequence the predicate letters can be interpreted according to any predicate (for otherwise we would not be generalizing over all the relevant arguments). Now, the generality absolutist wants to make use of such a notion when  $\forall$  is understood as being absolutely general. But then the following argument needs to be accounted for:

1. For everything  $o$ ,  $I(F)$  is an interpretation under which  $P$  applies to  $o$  if and only if  $o$  Fs.

According to the generality absolutist maximally liberal understanding of thing, even an interpretation such as  $I(F)$  counts as a thing. That is, the generality absolutist cannot say that he is only quantifying over things of a certain kind. He is quantifying over all things of all kinds and an interpretation certainly is a thing, regardless of its kind. Now define a verb 'R' in the following way:

2. For everything  $o$ ,  $o$  Rs if and only if  $o$  is not an interpretation under which  $P$  applies to  $o$ .

Now simply put 'R' for 'F' in (1) and apply (2):

3. For everything  $o$ ,  $I(R)$  is an interpretation under which  $P$  applies to  $o$  if and only if  $o$  is not an interpretation under which  $P$  applies to  $o$ .

Since 'everything' is unrestricted,  $o$  can be  $I(R)$  itself. But then (3) implies,

4.  $I(R)$  is an interpretation under which  $P$  applies to  $I(R)$  if and only if  $I(R)$  is not an interpretation under which  $P$  applies to  $I(R)$ .

(4) is obviously contradictory, so the generality absolutist needs to provide an account of what has gone wrong in the previous argument.

To deal with this, Rayo and Williamson (2003) suggest the adoption of the notion of *second-order interpretation* put forward by Rayo and Uzquiano (1999) and according to which second-order quantifiers are to be seen as plural quantifiers and second-order

variables as plural variables. One difficulty with Boolos (1985) treatment of second-order quantification is that it takes as primitive a satisfaction predicate that applies to first-level predicates and that as such can't be explained in terms of plural quantifiers. To free the notion of *second-order interpretation* from this difficulty, the generality absolutist might try to extend Boolos's approach to third and higher-order quantification. Rayo (2006), for instance, provides a semantic account of second-level predicates in terms of super plural quantification. However, also according to Rayo, adoption of this extension leaves us facing the following trilemma: we must choose between giving up absolutely general quantification, settling for the view that adequate semantic theorizing about certain languages is essentially beyond our reach, and countenancing an open-ended hierarchy of languages of ever ascending logical type.

In the present context, the first horn is to be immediately discarded. Adoption of the second horn seems to force the generality absolutist into saying that he is not able to make his position completely explicit. Rayo, however, claims that the third horn is an option available for the generality absolutist. We argue that this is not so and that consequently the generality absolutist can only adopt the second horn.

#### References:

Boolos, G., (1985a). Nominalist Platonism. *Philosophical Review* 94, pp. 327–44. Reprinted in Boolos (1998a).

Williamson, T., 2003. Everything. *Philosophical Perspectives* 17(1), pp. 415-465.

Rayo, A., and Uzquiano, G., 1999. Toward a Theory of Second-Order Consequence. *The Notre Dame Journal of Formal Logic* 40, pp. 315–32.

Rayo, A. and Williamson, T., 2003. A Completeness Theorem for Unrestricted First-Order Languages. In Beall, J.C., ed. 2003. *Liars and Heaps*. Oxford: Oxford University Press, Ch. 15.

Rayo, A., 2006. Beyond Plurals. In: Rayo, A. and Uzquiano, G., eds. 2006. *Absolute Generality*. Oxford: Oxford University Press, Ch. 9.

Williamson, T., 2003. Everything. *Philosophical Perspectives* 17(1), pp. 415-465.

## How numbers might be given to us

Mario Santos Sousa  
[msansou@gmail.com](mailto:msansou@gmail.com)

Dept. of Philosophy,  
University College London

Depto. de Lingüística, Lógica y Filosofía de la Ciencia,  
Universidad Autónoma de Madrid

Numbers are not the sorts of things we would expect to encounter. They cannot be seen, smelled, touched, tasted or heard. They do not leave traces from which their presence might be inferred (in contrast to electrons, say). Numbers are abstract entities, non-spatiotemporal and causally inert. Hence, they can neither impinge on our senses nor be detected by any instrumental means. This, however, seems to place them beyond our cognitive reach. For how could we possibly know about numbers, let alone form any reliable beliefs about them, if they cannot causally interact with anything in the physical world, including us?

The problem arises out of conflicting assumptions about the nature of arithmetic, on the one hand, and the nature of human cognition, on the other. If knowledge requires some sort of causal connection between the cognitive agent and the objects of knowledge, then knowledge of numerical abstracta does not seem to be possible. This raises an epistemological challenge. The following strategies suggest themselves:

- a. changing the description of the objects of knowledge in order to make them more amenable to human conditions on knowledge, or
- b. revising the conditions on knowledge to make them more amenable to our knowing about such (otherwise inaccessible) objects.

I will approach the matter from both directions and try to meet the challenge halfway, as it were. I will also simplify matters by considering how finite cardinal numbers might be given to us. Cardinals, if such they be, are anything but concrete. So how do we come to know about them?

The first thing we need to ask is, what are cardinal numbers? Cardinal numbers typically feature in response to ‘How many?’ questions, such as ‘How many coins are in my right pocket?’ The answer to this question gives the cardinal measure of the collection of coins in my right pocket, and that is much a property of the coins, considered collectively, as their aggregate value or their average weight. Hence, cardinals may be naturally thought of as properties of collections (if we allow for collectives of one and zero items). If we accept this, then we can make some progress towards explaining how we could possibly know about them: via their concrete instantiations. An account of how numbers might be given to us is not impossible if we take them to be properties. Still, this may be thought to be problematic. For what is the relationship between an abstract property and its particular instances? This of course is a problem, but it is a general problem that arises in many other areas of inquiry. So let’s put it this way: if cardinal properties exist, then we can have knowledge of (some of) them through their particular instances.

I will help myself to recent research in cognitive science, which I believe allows us to give a naturalistic account of such knowledge. The idea is that we start with an empirical hypothesis concerning how our numerical beliefs are formed and then provide an account that

shows how beliefs formed in that way amount to knowledge.

According to the available evidence, humans (both infants and adults), as well as other non-human species, have a “number sense”, which comprises a sense of small cardinal size and a rough sense of large cardinal size. This capacity is relatively abstract: it is not tied to any specific sensory modality, such as vision, but operates across different modalities (auditory, tactual, etc.). Experiments are designed to control for various non-numerical confounds that might correlate with number, such as duration for sequentially presented items and total surface area for static spatial displays. This rules out explanations of success in terms of non-numerical properties. The existing results have been widely established (and, thus, can be safely accepted) despite remaining controversies about the particular mechanisms that underlie these abilities.

Yet our sense of number is fairly narrow: we only discriminate small collections of objects up to a certain threshold (of about four) and approximate larger numerical quantities (failing to capture numerical differences below a certain ratio). Given these limitations, I will explain how, in my view, large cardinalities might be accessible to us by considering the role of numeral systems in enabling us to move beyond what our innate capacities can achieve.

We use all sorts of artefacts to supplement our natural senses: while our senses give us initial access to certain properties of the world, measuring instruments extend that access to other regions of the same property. Similarly, I propose to conceive of numeral systems as measuring instruments that allow us to detect large cardinalities, which we are not naturally equipped to discriminate. These epistemic devices initially build on our innate numerical capacities and are justified on the basis of sensation. Our counting experience enables us to map exact number sense representations of small cardinal size to an initial fragment of the numeral sequence in the presence of relevant stimuli, and then build further associations by mapping operations on collections and relations between numerals, until the relevant number sense representation gets activated when presented with the corresponding numeral. Then we can start reasoning about (small) numbers in terms of numerals, even in the absence of relevant stimuli. Beyond this stage, we let numerals do (some of) the thinking for us. We let them sharpen and even correct sensation as long as they bear sufficient conformity with our rough number sense representations of large cardinal size. Eventually, they extend our cognitive reach into cardinal regions that would otherwise be inaccessible to us, so that we start using numeral systems as a means of acquiring evidence to confirm that the relevant cardinal property has been instantiated. Numerals thus enable us to reason about cardinal numbers, any cardinal numbers, even when they are not perceptually given to us.

## **Two Forms of Dualism**

David Sosa

University of Texas at Austin

david\_sosa@austin.utexas.edu

I distinguish two sorts of motivation for dualism. One motivation is driven by the distinctive character of conscious phenomenology. The other is driven by the special character of normativity: is rationality an even "harder" problem than consciousness? There is no dramatic climax in which I show that these two dualist currents have a common source; in fact I think they are relatively independent.



## Non-propositional content, assertion and same-saying

Isidora Stojanovic

Institut Jean-Nicod, 29 rue d'Ulm, 75010 PARIS, isidora.stojanovic@ens.fr

In this talk, we argue that cases in which sentence content is a proposition are exceptions rather than the rule, and that this is true not just in general, but also for the subclass of *assertive* sentences. What is more, we focus not on those sentences that involve special lexical items or constructions, such as epistemic modals or dangling adverbs, which are admittedly problematic for propositionalist approaches. Rather, we argue that even in the case of most plain and simple sentences, like “I am hungry” or “She is late”, the content falls short of being a proposition.

The talk is divided in two parts. In the first part, we present our account of sentence content, in general albeit rough lines. For the sake of exposition, we will present it as a generalization of a fairly well-known account of *attitudinal* content put forward in Lewis (1979). In the second part, we motivate our account by arguing that it fares better than the more traditional propositionalist accounts on (at least) two scores: (i) it provides a better account of the same-saying relation (that is to say, of the conditions under which we are inclined to hear and report two speakers as “saying the same thing”); and (ii) it coheres with some plausible assumptions regarding the syntax-semantics-pragmatics interface.

### Short summary of Part I: applying Lewis' account of attitudes to assertion

Lewis (1979) proposes an account of beliefs in terms of self-ascription of properties (the account is primarily intended for *de se* attitudes, but generalizes to all other attitudes). For example, the content of Ann's belief that she is hungry, which she might express by saying “I am hungry”, is, according to Lewis, just the property of being hungry, a property that Ann ascribes to herself. Lewis never held, though, that the content of *assertion* should be analyzed along the same lines, and the dominant view is that if Ann says “I am hungry,” the asserted content is the proposition that Ann is hungry (at a given time). In this talk, we will extend Lewis's proposal for *de se* attitudes to assertion. The content of Ann's utterance of “I am hungry”, we suggest, is (equivalent to) the property of being hungry, and is a property that Ann is asserting of herself. In greater generality, however, we will depart from Lewis in two respects: (i) while Lewis' contents are always properties, i.e. map triples (agent, time, world) to truth values, our contents are mappings from possibly richer sequences, involving e.g. a pair of individuals (rather than a single agent), a location, a scale, etc. (ii) while Lewis generalizes the account to the 3<sup>rd</sup> person (or *de re*) case by construing it as a special kind of the *de se* case, viz. one in which the agent self-ascribes the property of inhabiting a world in which a certain proposition is true, we see the content as being the same regardless whether the sentence is in the 1<sup>st</sup> person or in the 3<sup>rd</sup> person; the difference is that in the former case, the speaker *self*-asserts

this content while in the latter, she asserts it *of* the person or the object that she is talking about. Both points may be illustrated by considering “She is late”. Suppose I say this while pointing to Ann. The asserted content is, then, the property of being late, and is a property that I assert of, or about, Ann. What is more, being late is a *relational* property: to be late is to be late *for* something. This, in turn, gives us for a content that maps quadruples (individual, event, time, world) to truth values; a content that, let's suppose, I assert of Ann and of the concert tonight.

#### Short summary of Part II: motivations

Our primary motivation for applying Lewis' account of *de se* belief to assertion is the fact that if two speakers say “I am hungry”, there is a strong intuition that there are saying the same thing – for, each is saying that he or she is hungry. The dominant, propositionalist views usually explain this intuition by pointing out that such speakers are, after all, uttering the same words, and in this sense may be said to have “said the same thing.” We argue that this response doesn't work, because using the same sentence is neither a necessary nor a sufficient condition for same-saying. Thus suppose that, at the ENS, Ann says “I study here”, and suppose that Bill says “I study at the ENS”. If one replies to Bill, “That's what Ann said, too”, the report is ambiguous between reporting Ann as having said that Bill studies at the ENS versus that *she herself* studies there. There is thus an important sense in which Bill and Ann said the same thing – they said that they study, each, at a certain place, viz. ENS. On the other hand, using one and the same sentence does not warrant same-saying. Thus suppose that Ann, pointing at Bill, says “He is crazy”, and that Carol utters the same sentence pointing at Dan. If one replies to Carol, “That's what Ann said, too”, the reply, as it stands (and without special context) is not ambiguous, but downright false. There is a striking asymmetry between the first person pronoun 'I' and the rest, and we will show how our Lewisian account of assertion can explain the asymmetry.

These examples illustrate the kind of linguistic evidence coming from same-saying reports that we use to motivate our account. Our account also handles immediately and without difficulty the so-called “non-sentential” cases that have received some attention lately (partly due to Stainton (2006)). Additional motivation for our account comes from more theoretical considerations regarding the notion of semantic content. In particular, we show that the proposed notion fits well with the picture of the syntax semantics interface as outlined e.g. in Stanley (2000). We also think that it provides a plausible picture of the semantics/pragmatics interface, similar to the one outlined e.g. in Carston (2008).

#### Cited works

Carston, R. (2008): "Linguistic communication and the semantics/pragmatics distinction", *Synthese* 165: 321-346.

Lewis, D. (1979): “Attitudes *de dicto* and *de se*”. *Philosophical Review* 88: 513-543.

Stainton, R. (2006): *Words and Thoughts*. Oxford: OUP.

Stanley, J. (2000): "Context and logical form". *Linguistics and Philosophy* 23.

## What we do when we judge

Josefa Toribio  
ICREA / UAB Department of Philosophy  
Facultat de Filosofia i Lletres. Building B, 08193 Bellaterra, Barcelona.  
jtoribio@icrea.cat

There are two ways of denying doxastic voluntarism regarding judgments:

- *Involuntarism*: Judging is an unconscious, involuntary, type of mental episode and hence not an action at all [O'Shaughnessy, Strawson].
- *Standard non-voluntarism*: Judging is neither voluntary nor involuntary: it is a non-voluntary type of mental action. In judging, we exercise a complex of conceptual abilities governed by truth-conducive norms [Hieronymi, McDowell, O'Brien].

Much of what there is at issue in settling the question of whether judging is a type of mental action or not relates to how we are to understand epistemic responsibility. The standard non-voluntarist claims that our responsibility consists precisely of our being answerable to relevant rational considerations, while the involuntarist claims that those relevant rational considerations only shape the process of inquiry. But there is nothing agential in the judgments we end up making as a result of applying such rational considerations.

In this paper, I develop a view about the nature of judging in which I aim to reconcile the thesis that judging is a type of mental action and the claim that it is a non-voluntary action in an as yet unrecognized way. My view captures, on the one hand, the important role the involuntarist assigns to spontaneity in judging, but such spontaneity is, I contend, constitutive of agency and thus supports the idea that judging is a type of mental action. I concur, on the other hand, with the non-voluntarist in viewing judging as a type of mental action, but the explanation of why this is the case, I argue, cannot be just that, in judging, we are answerable to rational considerations. Being thus answerable does not exhaust agency.

Judging is non-voluntary, I contend, not only because it is governed by epistemic reasons, but also, and importantly, because it is partially constituted by the exercise of a non-reason-governed, unconscious, skill. This skill consists of a pre-reflective, unconscious, pattern-recognition ability to spot the kind of situations in which the rationally relevant reflective abilities, which are also partially constitutive of the mental act of judging, should be deployed. We are responsible for our judgments, I conclude, because, in the pre-reflective way in which we spot those situations that demand the deployment of our reflective capacities, we reveal the kind of epistemic agents we are.



## Time Travel in a Branching Universe

Giuliano Torrenco  
University of Torino  
[giuliano.torrenco@unito.it](mailto:giuliano.torrenco@unito.it)

In recent articles Kristie Miller (2006, 2008) has argued that the Open Future hypothesis is incompatible with time travel. This incompatibility is not logical. Rather, endorsing both theses would force us to an unacceptable view of causality, according to which it can be indeterminate at a time  $t$  what are the causes of certain events that happens at  $t$ . For instance, if at a present moment  $t$  a person P pops up out of thin air and sincerely describes herself as a time traveller from the future, given that it is indeterminate at  $t$  which future branches will be connected to  $t$  (one in which P enters a time machine, or one in which she does not), it is indeterminate whether P's arrival has been caused by P's departure from the future or not (and thus whether P is a time-traveller or not). Given certain plausible constraints on causality – in particular, that the nature of the cause is always at least as determinate as the nature of the effect – this situation leads us to reject either the hypothesis that time travel is possible or that the future is open. Since the possibility of time travel seems to be a lively option in the present debate in physics, Miller argues that it is the idea of the open future that has to go.

Although I take Miller's conclusion to be on the right track, her argument seems to rest on a certain interpretation of the Open Future thesis, which is questionable. There are two intuitions behind the idea that the future is open: (i) all future alternatives are metaphysically on a par, and (ii) future tensed claims do not have, with respect to the present, a determinate truth-value. These intuitions have been formalized through the branching theory of time (Belnap et al 2001, MacFarlane 2003, 2008) and a non-bivalent logics (for future tensed propositions), based on such branching models. The branches represent nomologically possible alternatives, and are all temporally connected to the present (they are in the same world), although are not temporally connected to each other. Note that the tenet that *time has a branching topology* is weaker than *the branching theory of time*, which comprises the further tenet that bivalence fails for future tensed claims. In particular, if the having a determinate truth-value for a future tensed claim does not boil down to having the same truth-value with respect to every branch, the two intuitions may diverge. We can think at *the* future as a "thin red line" that singles out one branch among the others, as the way the world will be.

The Thin Red Line view has been criticized as not being a genuine alternative to the block view of time (Belnap & Green 1994, MacFarlane 2008), but the arguments against it are not compelling. Indeed, the difference between the block view and the red line view is quite clear: in the red line view, differently from the block view, all possible futures are temporally connected to the present (they all share the same world), although, differently from the branch view, it is determined, in the present already, what branch will be the actual future. The central claim of the paper will be that if time in our universe has a branching topology then we have an open future, and that time travel in a

branching Universe is possible, in so far as the thin red line view, and not the branching theory of time, is true.

Miller identifies the Open Future thesis with the branching view of time. However, it is not the mere topological structure of the branching model that warrants the thesis that it is not determined, at the present time  $t$ , what branch among the nomological possible ones will be actual. The thin-red line view combines a tree-like topology with the idea that there is a single actual future (the red line). The red line model is topologically indistinguishable from the branching model (they are both tree-like models), but does not suffer from the problems that Miller raises with respect to time travel. Reconsider the case of a person P popping out of thin air in the present possessing memories of being a time traveller from the future. If one of the future branches in which P gets in a time machine and travels back in time turns out to be how things will actually go, then P is a time traveller, otherwise she is not. In the branching model (with no red line), while it is determined that the arrival takes place at the time of its occurrence  $t$ , at  $t$  it is not determined whether its cause (the departure) will take place or not. In the red line model, it is determined (at  $t$ ) whether the cause will actually take place, and thus whether P is really a time traveller or not: we are not compelled to endorse any weird notion of causality in order to have both future branches and time travel. Note that in this view it can be determined at  $t$  that event C will take place at a later time  $t'$ , even if C is *not* on every future branch (at  $t'$ ), and thus if it is *not* nomologically determined (i.e. determined by the past states of the universe together with the physical laws). A fortiori, the thin red line view is compatible with the thesis that it is not metaphysically determined which branch will be the actual one. In the conclusion, I defend the idea that the thin red line view is indeed compatible with the thesis that the future is open, once we give a more precise characterization to the openness intuition, and thus Miller's argument is better described as an argument against the branching view of time, rather than against the Open Future thesis.

## References

- Belnap N., and Green M. (1994), "Indeterminism and the Thin Red Line", *Philosophical Perspectives*, 8: 365-88.
- N. Belnap, M. Perlo and M. Xu (2001), *Facing the future*, Oxford University Press, Oxford.
- J. MacFarlane (2003), "Future contingents and relative truth". *Philosophical Quarterly* 53: 321-336.
- J. MacFarlane (2008), "Truth in the Garden of Forking Paths", in M. Garcia-Carpintero and M. Kölbel (eds.), *Relative truth*, Oxford University Press, Oxford: 81-102.
- K. Miller (2006), "Time Travel and the Open Future". *Disputatio* 19(1) 197-206.
- K. Miller (2008), "Backwards causation, time and the open future". *Metaphysica* 9(2): 173-191.

## Epistemic privilege and our knowledge of artifacts

Jesús Vega Encabo  
Universidad Autónoma de Madrid  
Dept. Lingüística General, Lógica y Filosofía de la Ciencia  
Ctra. Colmenar, km. 15,4 – 29049 Cantoblanco (Madrid)  
E-mail: [jesus.vega@uam.es](mailto:jesus.vega@uam.es)

In a recent series of papers, A. Thomasson has argued that we have a certain epistemic privilege regarding the knowledge of artifacts and other cultural objects (including institutions). What kind of privilege is this? Where does this privilege derive from?

Thomasson argues from a certain metaphysical principle on the nature of artifacts to support her thesis. A certain *dependence principle* for artifacts states that the existence of an artifact cannot be independent of our knowing something about it, of having a correct substantive concept about what the K is. Artifacts are ontologically dependent on our beliefs and concepts.

The fact that artifacts are metaphysically dependent on the concepts and beliefs of the makers explains why we have a certain guarantee against massive error and against universal ignorance. We could not be completely mistaken or have a complete ignorance about what the artifact is. This contrasts with our knowledge of natural kinds. Regarding natural kinds experts are in a better position than me for knowing what the relevant features of the members of the kind are. And it is so because they have a certain familiarity with the members of the kind, a familiarity that derives from observation and inference and that I do not have. In this case, the notion of epistemic privilege could be understood as simply being in a better position to get true and reliable information about a subject matter. The expert has this kind of epistemic privilege over the layman.

But this cannot be the notion of epistemic privilege that we seem to apply to our knowledge of artifacts. Makers are not experts in this sense. There could be experts that derive a better epistemic position on artifacts just by having an “extensive interaction” with them (Kornblith). But this does not explain the epistemic privilege of the maker in the relevant sense that would account for this guarantee against massive error and universal ignorance. Moreover, it is also true that merely to insist on the idea that this privilege derives from the fact that intentions are *constitutive* of the nature of the artifact K (Thomasson) does not explain too much either, because this thesis could be understood as just stating that if the artifact exists, then someone (the maker/designer) must have non-mistaken knowledge about the artifact and its *real* nature. But this does not account for the privilege of the kind of knowledge she has and that apparently serves to guide the production of the object and *constitute* the object as such.

In this paper, I examine several alternatives to account for the maker’s knowledge of artifacts and the type of epistemic privilege we ascribe to her. I take as my starting point a contrast already clearly stated in Maimonides between the maker’s knowledge and the knowledge others could have of the same object. Consider the following quotation by Maimonides:

“THERE is a great difference between the knowledge which the producer of a thing possesses concerning it, and the knowledge which other persons possess concerning the same thing. Suppose a thing is produced in accordance with the knowledge of the producer, the producer was then guided by his knowledge in the act of producing the thing. Other people, however, who examine this work and acquire a knowledge of the

whole of it, depend for that knowledge on the work itself...” (Maimonides, *Guide for the perplexed*, part III, chap. 21).

In the terms proposed by Maimonides, the maker’s knowledge seem to be a kind a authoritative knowledge that only the maker can have of the products of her activities. Besides maker’s knowledge is not observationally acquired. It is knowledge that guides the actions that are then “produced according to the knowledge”. But in virtue of what the maker knows whatever she knows about the particular artifact she is making? And what guarantees that she is not wrong or universally ignorant about it? And is the kind of privilege that she seems to exhibit accountable in terms of the authoritative knowledge she has of her own intentions and actions? These are the questions that will be addressed in the paper. And I will suggest first that it is not possible to explain maker’s knowledge in terms of the knowledge of intentions and second that making by itself cannot yield authoritative knowledge of the product. We cannot bring into existence an artifact just by *fiat*, in the way that God could be the maker of the laws of the universe (as Maimonides seems to suggest).

I will argue that Thomasson cannot adequately explain the knowledge the maker has of the nature of the artifacts and her epistemic privilege. I will try to show that Thomasson’s model depends on the assumption that the nature of the artifact can be fully established by stipulation, on the model of our creation of fictional entities. It is true, I cannot bring into existence a chip merely through performing linguistic acts (as I could do in the case of fictional entities), but what I can do is to determine what a certain K is by stipulating that it has to comply with certain criteria (provided that the object finally exists). That is why, following Thomasson, I cannot be mistaken or be ignorant about what a K is. Thus Thomasson is assuming that the relevant features that characterize the artifact are stipulatively determined by the maker, provided that they become successfully realized. So in creating the artifact the maker knows because she establishes by *fiat* what the nature of the artifact is. In a sense, knowing precedes what is made and is independent of the making itself.

But there is also a long tradition where the maker’s knowledge is constituted by the very making itself. That means that it is the making that gives some knowledge of what is made. How to spell out this idea? Here we confront an obvious dilemma: if we take that making is an intentional and purposeful activity, then it assumes that the intentional maker has some previous knowledge (Mackie); but if we postulate a particular kind of “making” that constitutes by itself a certain kind of knowledge, then it seems very implausible that *our* making and knowing of artifacts (maybe in contrast with the kind of intuitive knowledge expressed by God), fits this model.

Finally, I will suggest that if making provides by itself knowledge of what is made is because of the kind of understanding that the maker exhibits and acquires at once in the making of an artifact. In a sense, our knowledge of what we are making is also dependent on the interaction with the object itself, and with other objects, and on our control of causal patterns in the world. This does not mean that the knowledge of what she is doing when she is doing it is observational, insofar as this (practical) knowledge provides a description following which what happens constitutes the execution of the intention, as Anscombe remarks. But this does not constitute by itself our knowledge of the nature of what we are making through our actions.

## Fregean and Anti-individualist Concept Possession: A Proposal of Compatibility

Víctor M. Verdejo

Dept. Lingüística General, UAM, Ctra. Colmenar km. 15,6, Cantoblanco 28049 Madrid  
victor.verdejo@uab.cat

The following are widely held (Fregean and anti-individualist) theses in the theory of concepts: 1) concepts are individuated by considerations of cognitive value of rational subjects; 2) concepts are individuated at least in part by environmental conditions (physical or social) external to the subject. Although widespread, these theses can be seen as exhibiting a not obviously surmountable tension, as has been shown in discussion regarding self-knowledge of thought contents for the last twenty years (see Brown 2004 for a comprehensive analysis). However, a not so attended aspect of that tension appears in relation to concept understanding or *concept possession* (Wikforss 2008). Arguably, concept possession is *the* central issue in the debate between Fregeanism and anti-individualism. First, (neo-) Fregean approaches result in accounts that specify concept-individuating possession conditions (Peacocke 1992). Furthermore, the main conclusion of Burge's 1979 seminal paper can be taken to be about "how *having* certain thoughts constitutively depends on relations to the environment" (Burge 2006, p. 156, emphasis his). My concern presently shall thus be in the analysis of the compatibility of 1) and 2) as regards the following theses:

Fregean possession conditions (FPC for short): Concept-individuating possession conditions are determined by considerations of cognitive value of rational subjects.

Anti-individualist PC (APC for short): Concept-individuating possession conditions are determined, at least in part, by environmental (social or physical) conditions external to the subject.

In the literature one finds several ways of providing a compatibility reading of FPC and APC. Here I focus on three:

A) The staunch anti-individualist view consists in claiming that cognitive value must itself be socially articulated and inaccessible to the subject (Goldberg 2008). A thought experiment can show that this position accounts poorly for cases in which two subjects have different cognitive stances associated to a concept –and hence arguably different concepts– but which are in the very same deferential setting.

B) The staunch individualist position (Bilgrami forthcoming) contends that, even if environmental conditions have some role, cognitive value stands for knowledge of contents transparent to the subject. An intuitive analysis makes explicit that such position leads to the unacceptable claim that for every deviant interpretation  $D_1, D_2, \dots, D_n$  of a given standard concept  $C$  that a subject may have, we must acknowledge the subject's possession of a distinctive concept  $C_1, C_2, \dots, C_n$ , corresponding to each deviant interpretation. Auto-satisfying possession conditions (Higginbotham 1998) and the non-publicity of concepts follows straightforwardly (Fodor 1998).

C) Peacocke's middle way has been expressed in some of his writings (e.g., Peacocke 1992, pp. 29-30 and 1997, p. 13) consists in claiming that core cases of concept possession must be only cases of full understanding or mastery of the concept. This means that APC is only OK for cases of mere attribution of the concept, but not in paradigmatic accounts of possession conditions. I argue that Peacocke's suggested approach is deficient because: a) it distorts the classical analysis of differing cognitive values; b) it turns on possession conditions of concepts the great majority of people lacks; c) can be interpreted as boiling down to the staunch individualist position.

I propose to distinguish two aspects in the possession conditions of a concept C: i) what subjects must know for possession of C; ii) what condition must a subject meet for qualifying as having the required kind of knowledge for possession of C. i) but not ii), I argue, involves anti-individualist components such as a) what worldly reality the knowledge is knowledge of; and not so attended in the literature, b) what conventional degree of knowledge is sufficient for possession of C. By contrast, ii) must be understood in terms of individualistically determined knowledge (be this knowledge considered in terms of inferential role, linguistic competence or informational states). This sheds light and articulates a distinctive notion of cognitive value as an anti-individualistically determined kind of knowledge.

#### References

- Bilgrami, A. (Forthcoming). Why Meaning Intentions are Degenerate. In Wright C. (Ed.), *Volume of Essays*. Oxford: Oxford University Press.
- Brown, J. (2004). *Anti-individualism and Knowledge*. Cambridge, MA: MIT Press.
- Burge, T. (1979). Individualism and the Mental. Reprinted in *Foundations of Mind*, T. Burge, Ed. (Oxford: OUP, 2007), 100-150.
- Burge, T. (2006). Postscript to "Individualism and the Mental". In T. Burge (Ed.), *Foundations of Mind*. Oxford: Oxford University Press.
- Fodor, J.A. (1998). *Concepts. Where Cognitive Science Went Wrong*. Oxford: Oxford University Press.
- Goldberg, S. (2008). Must Differences in Cognitive Value be Transparent?, *Erkenntnis*, 69, 165-187.
- Higginbotham, J. (1998). Conceptual Competence. *Philosophical Issues*, 9, 149-162.
- Peacocke, C. (1992). *A Study of Concepts*. Cambridge, MA: MIT Press.
- Peacocke, C. (1997). Concepts Without Words. In R. Hecke (Ed.), *Language, Thought, and Logic: Essays in Honour of Michael Dummett*. Oxford: Oxford University Press, 1-33.
- Wikforss, A. (2008). Self-knowledge and Knowledge of Content. *Canadian Journal of Philosophy*, 38(3), 399-424.

## **Emergentism: complex systems and selection processes**

Agustín Vicente

Ikerbasque/ UPV-EHU

agustin\_vicente@ehu.es

The last years have seen a renewed interest in the emergentism issue. The contemporary debate, in contrast with past times, has to do not so much with the mind-body problem as with the relationship between the physical and other domains, mostly with the biological domain. Philosophers discuss whether there are proper (i.e. non-reducible) non-physical properties or whether, as reductive physicalists claim, everything is reducible to the physical. One of the main sources of this renewed interest is the study of complex and, in general, far-from-equilibrium (FFE) self-preserving systems. FFE systems are intriguing from a philosophical perspective because it seems as though knowledge of constituents and of their causal powers does not suffice to predict the behavior of the whole system. Thus, FFE systems seem to fulfill one of the necessary conditions for an entity to be emergent, namely, that its causal powers are not predictable from the causal powers of basic physical properties. Take the case of a hurricane: it seems that the behavior of the particles forming part of it cannot be predicted from the knowledge of their fundamental physical properties, basically because their behavior is constrained by their being parts of a whole.

The more complex a FFE system is, the more we are tempted to see emergent properties in it. However, intriguing and perhaps impressing as self-maintaining FFE systems may seem, reductionists are not convinced. It's one thing to say that we are unable to predict a hurricane, a traffic-jam or the appearance of life; it's quite another to claim that they are unpredictable *in principle*. We simply cannot know whether enough knowledge would enable us to predict the formation and duration of a hurricane. Moreover, where can the new powers stem from? If everything is made of the same fundamental entities, how can it be that the behavior of all things is not due ultimately to the behavior of these fundamental entities? Emergentists do not have a clear answer, and the dispute between emergentists and reductionists has no end in sight.

Now, I think it is possible to show that there are certain causal explanations that cannot be substituted by physicalist explanations even as a matter of principle. However, I also think that much of current emergentism has misfired by being focused on the interpretation of FFE self-maintaining systems. If we want to find emergent properties, we should not look at the processes that lie at the heart of complexity, but at selection processes. Typically, a process of selection (or recruitment) involves that a mechanism of selection "chooses" an entity due to its instantiating one interesting property. Mechanisms of selection, however, are not constrained to choosing entities only on the basis of their physical or fundamental properties. Their target can be, and usually are, properties that, on any account of physicalism, would not count as physical properties, such as functional or relational properties. My claim, thus, will be that in selection processes functional or relational properties can be causally efficacious, which means that properties other than physical properties can have causal powers; even more, they can have "downward" causal powers.

The way the argument will be developed is: in the first section I will introduce and argue for the idea that explaining what goes on in selection processes involves resorting to non-physical emergent properties. When an entity is selected, the cause of its being selected is its instantiating a particular property. Generally, the properties that

are involved in selection processes are not physical fundamental properties, but, typically, functional or relational properties. Thus, hearts were not selected because they had certain physical or categorical properties, but because they were blood pumps. And an organism that has a better orientation sense than the rest of its co-specifics endures and has a higher rate of reproduction precisely because of that, i.e. precisely because it stands in a certain relation to its co-specifics. The same can be said about many other non-Darwinian selection processes: if we want to explain why a certain antelope died in the hands of a lion, we have to mention, e.g., that it was the weakest of its herd.

In the second section, I will argue, first, that the information captured by these explanations that resort to non-physical properties cannot be captured in any other way, and, second, that, because of this, such explanations have to be considered as strictly causal, i.e. explanations that mention causes directly. The first issue can be put in terms of modal information: Take the relational property *being the weakest of the herd*. This property does seem to draw a sharp division in the sphere of possible worlds: on the one hand, we have the worlds where the antelope dies (those worlds where the antelope is the weakest of its herd); and, on the other, the worlds where it survives (those worlds where it is not the weakest of the herd). This kind of modal information cannot be captured by any physicalist-friendly explanation. Now, secondly: in the light of this, is there any reason to deny that the causal explanations we are considering are strictly causal? My claim will be that, given that they cannot be substituted by any other explanation that captures all the causal information they provide, these explanations are not threatened by any exclusion problem whatever. And an explanation not threatened by an exclusion problem can be considered a strictly causal explanation.

Finally, I will face and discuss the perennial problem of emergentist proposals, namely, how to conciliate their claims with the so-called principle of the causal closure of the physical world (CCP). I will argue that my proposal is compatible with the truth of the principle, though it certainly clashes with what seems to be an epistemic consequence of the principle, namely, that every physical effect has a sufficient (complete) physical explanation. Thus, I will be denying that all physical *explananda* have sufficient physical causal explanations, at least when 'sufficient' is understood as complete, or to put it more vividly, as completely complete. However, I don't find this result problematic. It may be that the CCP is a truth of physics, and accordingly, a truth that philosophers are not in a position to deny. But the claim that physics can fully explain all that happens in the physical world is not a truth of physics, and, in fact, it does not follow from the CCP either.

## La paradoja de Moore y la relación entre aserción y creencia<sup>1</sup>

Javier Vilanova Arias  
Universidad Complutense - vilanova@filos.ucm.es

En este trabajo pretendo defender la siguiente tesis: la función primaria de la aserción de *p* es la adopción de un compromiso con la verdad de *p* y no la expresión de la creencia de *p*. Un importante problema para esta tesis es el generado por la llamada “paradoja de Moore”, según la cual la aserción de la siguiente oración constituiría una contradicción:

(1) “*p* pero yo no creo *p*” (donde *p* es una oración asertiva cualquiera, por ejemplo: “el gato está sobre la alfombra”).

Según algunos autores el absurdo de la paradoja de Moore sólo se puede explicar por el hecho de que la afirmación de una oración *p* por parte de un individuo *X* implica en algún sentido (al menos si el individuo *X* está siendo honesto) que *X* cree que *p*. De ahí que si el individuo afirma que *p* y a continuación afirma que no cree que *p* se contradice a sí mismo. En esta línea se inscriben autores como Black 1952, para quien la creencia en *p* es una “presuposición” de la aserción de que *p*, o Searle 1979, para quien hay una “conexión interna” entre el acto de habla de afirmar que *p* y el estado psicológico de la creencia en *p* (p. 192), o Austin 1961, para quien el aseverar que *p* implica en un sentido especial, no lógico, “yo (quien asevera) cree que *p*”.

Para empezar, es importante darse cuenta que “afirmar que *p*” no implica “creer que *p*” en el sentido de implicación lógica ni, al menos de manera general, implicación analítica o semántica, como se desprende de la observación de casos en los que, al contrario de lo que ocurre con la paradoja de Moore, no asoma la mínima sombra de inconsistencia:

(2) *p* pero yo no creía que *p*,

(3) *p* pero Juan no cree que *p*.

Como es sabido, este hecho es utilizado por teóricos de los actos de habla para probar que es el acto de habla llevado a cabo por la preferencia de (1), y no (1) mismo, el que es inconsistente. De ahí se seguiría que la paradoja de Moore es el producto al menos en parte de las reglas o convenciones que rigen nuestros actos ilocucionarios, y no sólo de principios o propiedades generales de nuestras actitudes de creencia. Sin embargo, esta conclusión resulta, cuando menos, apresurada, ya que una oración como (1) no sólo resulta anómala en tanto que aserción, sino que también resulta anómala como contenido de una creencia. Es decir, nos resulta difícil concebir que una persona crea que *p* y al mismo tiempo crea que no cree que *p*. De otro modo, si:

(5) Afirmo que *p* y no creo que *p*,

es anómalo, no lo es menos (de hecho, como defenderé, lo es más):

(6) *X* cree que *p* y que no cree que *p*.

De facto, lo que repele nuestra intuición es lo mismo al oír (1) que (6): sentimos que nadie puede albergar a la vez la creencia de que *p* y de que no cree que *p*, como nadie puede albergar a la vez la creencia de que *p* y que no *p*. Pero si esto es así, entonces la “anomalía” de la paradoja de Moore no radica en la violación de reglas ilocucionarias,

---

<sup>1</sup> Este trabajo ha sido llevado a cabo en el marco de los proyectos de investigación FFI2008-03092 y FFI2009-08828/FISO financiados por el MCI.

sino de un principio doxástico, una regla de la lógica de la creencia, si se quiere. A saber:

(7) “X cree que p” implica que “X no cree que no cree que p”,

Por mucho que estas consideraciones “enturbien” la pretendida cogencia del argumento que Searle y otros defienden, todavía se puede derivar de la paradoja de Moore una evidencia indirecta de la conexión esencial entre afirmación y creencia. En efecto, si bien la anomalía detectada en la paradoja de Moore radica, en último término, en la violación de un principio doxástico, lo cierto es que dicha violación se efectúa a través de la preferencia de (1). Se sigue, pues, que de la preferencia de (1) por parte de X inferimos (6), que es contradictoria, y de ahí inferimos que (1) es contradictoria ella misma. Pero entonces de “X afirma p” estamos infiriendo “X cree p”.

Yo no niego que, en los casos en los que la preferencia de (1) nos resulte contradictoria, tal inferencia esté ocurriendo. Ahora bien, ¿qué es lo que hace de ella una “buena inferencia”?, y ¿es siempre una “buena inferencia”?. Aquí es donde nos encontramos dos respuestas muy distintas. Para Searle y compañía, la inferencia se apoya sobre una conexión “fuerte” entre afirmar y creer, una conexión dada por la propia naturaleza de nuestro lenguaje, por algún hecho semántico o una regla ilocucionaria que liga esencialmente aserción y expresión de creencias. Según la postura que yo defiendo la función principal de la aserción no es expresar creencias: mediante la afirmación de la oración p, el hablante se compromete con la verdad de p y sus consecuencias. Es verdad que, en tanto que agentes autónomos y libres, los seres humanos habitualmente adoptamos compromisos de verdad con oraciones que creemos, pero esto no es algo constitutivo del hecho de afirmar, y cabe esperar múltiples excepciones, ya que la autonomía agencial de la persona no es constante ni unilateral. Es de esperar pues, que aparezcan con relativa frecuencia situaciones en las que la adopción del compromiso de verdad no vaya acompañada de la creencia: precisamente situaciones en las que el hablante podría afirmar algo como (1), es decir, en que la situación descrita por la paradoja de Moore sea la situación real. Estas situaciones son “anómalas” en cierto sentido (se desvían del caso más frecuente o habitual) pero no son en absoluto “inconsistentes”, “absurdas” o “sinsentido”.

En el resto de la comunicación presento ejemplos de dichas situaciones: contextos argumentales, en los que un participante se ve obligado a aceptar una conclusión a pesar de su opinión adversa, contextos de negociación, en los cuales uno adopta el compromiso de actuar como si p fuera verdadera aunque no crea que p, y circunstancias de carácter institucional en las que uno hace una declaración oficial contraria a sus verdaderos sentimientos.

#### REFERENCIAS.

- J. L. Austin (1961): “The meaning of a word”, *Philosophical Papers*, Oxford Oxford U.P. (versión española (1989): *Ensayos Filosóficos*, Madrid: Alianza Editorial)
- M. Black (1952): “Saying and Disbelieving”, *Analysis* XIII, pp. 25-33.
- Hintikka, Jaako (1962): *Knowledge and Belief*, Ithaca: Cornell University Press.
- J. R. Searle (1979): *Expresión and Meaning. Studies in the Theory of Speech Acts*, Cambridge: Cambridge U. P.

# Minimal Expressivism

Neftalí Villanueva & María J. Frápolli

Departamento de Filosofía I, Universidad de Granada

[nef@ugr.es](mailto:nef@ugr.es)

[frapolli@ugr.es](mailto:frapolli@ugr.es)

## Introduction

The purpose of this paper is twofold. Our aim is, first, to outline a version of non-descriptivism, ‘minimal expressivism’, that leaves aside some of the long-standing problems that have been traditionally associated with expressivist views. Secondly, we will show how familiar expressivist’s results can be accommodated within this framework, through a particular interpretation of the role that the *expressive realm* has to play in a theory of meaning. Expressivist theories of meaning only face some of the classical problems attributed to this position when they undertake the task of offering an explanation concerning *why* the expressions they deal with have the meaning that they have. A position can nevertheless be dubbed ‘expressivist’ –in the minimal sense that we favor– simply by paying attention to the following key-features of the meaning of these expressions: they can be used as functions of propositions, and they are not used to describe how the world is.

Expressivism is a way to approach the meaning of a certain kind of expressions. A theory of meaning can be called ‘expressivist’ if it holds at least two of the following theses:

(1) Certain predicables do not take simple objects under their scope, but complexes of properties and their bearers, i.e., propositions. These predicables are ‘second-order’ predicables, or, as we prefer, *functions of propositions*. Not every second order predicable is a function of propositions, though, some of them are functions of functions. Examples of this latter kind are standard quantifiers, some uses of negation and conjunction, first-person operator (according to Anscombe-Wiliams' and Recanati's views) and others. This paper will focus on a particular subset of second-order predicables, those that can be used to produce propositions out of other propositions<sup>1</sup>. They are non truth-functional –the truth of the proposition that results from assigning specific propositions as its arguments is not a function of the

---

<sup>1</sup> Thesis (1) can be easily extended to cover predicables which arguments are predicables, and not complete propositions. It would suffice to modify (1) in the following sense;

(1') Certain predicates do not take single objects under their scope but n-adic predicables ( $0 \leq n$ , being 0-adic predicables propositions)

This modification would incorporate standard quantifiers, the first operson operator (in Anscombe, Williams, or Racanati's sense) and second order identity among the candidates for an expressivist treatment. Second order predicables would then become functions suitable to produce propositions out of predicables with any adicity.

truth-value of its constituents— and they are not extensional —embedded co-extensional expressions cannot be intersubstituted *salva veritate*. At least one of the items of the following list can be explained as an instance of this kind of propositional functions: Belief, knowledge, a priori, necessity, possibility, good, bad.

(2) These functions of propositions do not *describe* the way the world is.

(3) Expressions containing these functions of propositions lack truth conditions, even if they are syntactically correct —they are not ‘truth-apt’.

(4) These functions of propositions are used to express some attitude A towards a particular piece of content.

Expressivism is a thesis about the meaning of a particular kind of notions, and it is perfectly sound to be an expressivist about some higher-order notions and being a descriptivist about some others. For instance, logical expressivism, the position held by Wittgenstein and Brandom about logical constants, doesn't force anyone to be an expressivist about moral notions, and the other way around —a moral expressivist such as Stevenson is not necessarily committed to logical expressivism. We saw that most of the usual criticisms presented against expressivism were not aimed at theses (1) and (2), that spotted the presence in natural language of a special group of expressions of a non-descriptivist nature, but rather at theses (3) and (4). Thesis (3) posits that complex expressions containing counterparts of higher-order expressions lack truth conditions, but only (3') is a consequence of (2) —provided that their job is not to describe how the world is, second-order expressions do not modify the truth-conditions of the expressions falling under their scope. Some other objections concerned thesis (4), the specific characterization of the positive work that second-order expressions do. We think that (4) is not necessary for a theory of the meaning of higher-order expressions, since the meaning of these expressions is exhausted once their inferential potential is spelled out. This map of possible inferences would make apparent the relationships between some higher-order notions and others. *Positive* expressivist accounts of the meaning of these expressions are only worthy —from the point of view of the things that we as speakers can do with them— as long as they inspire hypotheses that might modify the inferential map. Here we have sketched a semantic view on a group of expressions particularly reluctant to traditional analyses. But this is only a first step. On similar lines, other sets of philosophically relevant notions admit an expressive treatment. Logical connectives, classical and non-classical quantifiers, first-person and identity, and the set of factive notions (necessity, knowledge and truth) will be illuminated from the point of view of their inferential behaviour once that we get rid of the classical descriptive approaches that have led most debates on higher-order notions to a disheartening dead-end.

## The myth of the inheritable gap

Neftalí Villanueva Fernández & Stavros Assimakopoulos

[nef@ugr.es](mailto:nef@ugr.es) [stavros@ugr.es](mailto:stavros@ugr.es)

Departamento de Filosofía I, Universidad de Granada

According to the received view on expressivism, the thesis that certain group of terms, characteristically ethical terms, do not *describe* how the world is, but *express* our attitudes towards a state of affairs, faces two groups of objections. The first group of criticisms has to do with objectivity: if ethical terms do not describe, how can we *know* that something is good or bad, correct or incorrect? The second set of criticisms allegedly pose a “logical challenge” to the expressivist proposal on the meaning of such terms (see e.g. Pendlebury 2010). It is this second objection that we are interested in in this paper. In a nutshell, if ethical terms do not describe, then complex expressions containing these terms are not “truth-apt” (see Lenman 2003); they are not suitable to express true or false propositions. How is then possible, the critic continues, to find these complex expressions containing ethical terms as arguments of truth-functional natural language counterparts of logical constants, which, by definition, can only receive truth-apt entities as arguments?

The aim of our talk is to show that the thesis that complex expressions properly containing non-descriptive constituents are not truth-apt crucially rests on an assumption that we call *the myth of the inheritable gap*, and which is neither theoretically motivated nor empirically adequate. We will not only show the genesis of this idea, but also the kind of counterexamples that a theorist willing to maintain such a thesis would have to face. The *myth of the inheritable gap* simply states that non-descriptive constituents are *infectious*. Whenever we have a non truth-conditional –in a sense to be specified below– gap in a sentence, the whole sentence cannot be used to express a complete proposition, something which can be declared as true or false. We think that this claim, as stated in its full generality, is false, and consequently we dub it “a myth”, but this by no means amounts to saying that it is never true that gaps can compromise the truth-conditional character of complex expressions. Meaningless expressions, and therefore truth-conditionally irrelevant, in subject position, for example, typically –if not always– render the whole complex as non truth-apt. Thus we cannot determine whether ‘D34F68Iâ is tall’ is true or false in a situation in which no contextual cues guide our interpretation of ‘D34F68Iâ’ so as to fix its deviant reference. The usual explanation of this phenomenon makes use of *compositionality*, the thesis that the meaning of a whole is a function of the meanings of its parts, and the way they are put together, and concludes that subsentential gaps have to be inherited by bigger structures, resulting in non truth-apt complex expressions, expressions that cannot be used to express true or false propositions. It will be defended that compositionality on its own is logically independent from the *myth of the inheritable gap*, whose origin is more likely to be found in a particular interpretation of the *doctrine of the unity of the proposition* (see. e. g. Linsky 1992). Our point will not be merely historical, though, as it will be argued that the arguments needed to support this particular version of the *doctrine of the unity of the proposition*, according to which the constituents that allow

the proposition to remain unite have to be of a descriptive nature, are completely independent from the traditional arguments supporting both the aforementioned doctrine and compositionality.

Finally, two sets of counterexamples will be presented in order to challenge the empirical adequacy of the myth of the inheritable gap. The first group will comprise truth-apt sentences containing non truth-conditional sentential operators, such as ‘sadly’, ‘frankly’, ‘unfortunately’, ‘confidentially’, etc. Such sentence adverbials have been traditionally treated as non-truth conditional, since they are taken to indicate the type of speech act performed by an utterance containing them or a propositional attitude towards it. Following research within Relevance Theory (see Ifantidou 2001) we will show that the argument of the inheritable gap is not only counterintuitive but can also become particularly shaky, especially since there are other cases of sentence adverbials, such as ‘allegedly’, ‘reportedly’, ‘obviously’, ‘clearly’ etc. which can be shown to directly affect the truth conditions of the sentence they are part of. The standard tests that are used to test the truth-conditionality of the sentences containing the adverbials involve embedding the sentence into the antecedent of a conditional and seeing whether it falls within the scope of ‘if’, or combining it with factive connectives such as ‘although’ and ‘because’ and seeing whether it falls within their scope too.

To conclude, a theorist willing to maintain some of the criticisms associated with the *logical challenge* should either motivate differently the thesis that truth-conditional gaps are *infectious*, and provide an alternative explanation for the kind of examples that we gave above, or take an alternative approach in order to support the arguments that depend on the Frege-Geach problem.

## References

- Ifantidou, E.(2001). *Evidentials and relevance*. Amsterdam; Philadelphia: J. Benjamins.
- Lenman, J. (2003). "Disciplined Syntacticism and Moral Expressivism." *Philosophy and Phenomenological Research* LXVI(1): 32-57.
- Pendlebury, M. (2010). "How to be a Normative Expressivist." *Philosophy and Phenomenological Research* LXXX(1): 182-207.
- Récanati, F. (2000). *Oratio obliqua, oratio recta :An essay on metarepresentation*. Cambridge, Mass.: MIT Press.
- Récanati, F. (2007). *Perspectival thought :A plea for (moderate) relativism*. Oxford: Oxford University Press.

## Relativism and Binding: The Case of Predicates of Personal Taste

Dan Zeman  
LOGOS - Universitat de Barcelona  
Montalegre 4-6, Barcelona 08001  
dan\_zeman@yahoo.com

The main concern of the paper is one argument that has recently surfaced in the debate between contextualism and relativism about predicates of personal taste (PPTs), argument that parallels the one found in the debate between truth-conditional semantics and truth-conditional pragmatics concerning the semantics of meteorological sentences such as “It is raining”.

In the first section I make clear how I’m going to use the terms “contextualism” and “relativism”, which given the multiplicity of positions and the different uses those two terms have received, seems a reasonable thing to start with. Thus, I use the term “contextualism” for any view in which context has a *content-determinative* role (in the terms of MacFarlane (2009)) and the term “relativism” for any view according to which context has a *circumstance-determinative* role.

Section 2 introduces PPTs (expressions such as “tasty”, “fun”, “disgusting” and their ilk) and presents the argument against relativism that Johnathan Schaffer (ms.) has recently put forward. Schaffer starts by noting that PPTs can have bound readings, as in

(1) Everyone got something tasty.

(1) has a reading according to which each person in the range of the quantifier “everyone” got something tasty *for that person*. The best explanation for this reading, and for making sense of the contrast between (1) and a sentence like “Everyone got something frozen” is to posit a hidden variable for the judge at the appropriate level of representation. Thus, Schaffer argues, (1) is to be rendered as  $\text{Everyone}_i$  got something tasty for  $x_i$ .

Section 3 draws the parallel between the argument above and the argument employed by Jason Stanley (2000) in order to argue in favor of truth-conditional semantic approaches to sentences like “It is raining” – an argument known as the Binding Argument. Stanley starts with the observation that

(2) Everytime John lights a cigarette, it is raining,

has a reading in which the location of rain is bound by the quantifier; that is, the truth-conditions for (2) are that for every time  $t$  that John lights a cigarette it is raining at  $t$  at the location in which John lights a cigarette at  $t$ . Since there is no binding without a bindable variable, Stanley claims that in the unembedded sentence “It is raining” a variable for locations is represented.

Section 4 surveys a number of answers to the Binding Argument that try to escape its conclusion. The first answer I consider is Peter Pagin’s (2005), which consists in the claim that the bound reading of (2) could be rendered by quantification over contexts in the meta-language instead of quantification over locations in the object-language. I point out problems with the view, based on some examples provided by Stanley (2005). Although I don’t fully agree with Stanley’s diagnosis of the cases, I nevertheless argue that Pagin’s attempted solution to the problem raised is unsatisfactory. The second answer to the Binding Argument that I survey is that given by Peter Lasersohn (2008): quantification over indices. In Lasersohn’s system, the quantifier “everyone” introduces both a pronomial element in the

syntax and a sentence-abstract-forming operator that binds the index in the meta-language. The main problem I find with Lasersohn's view concerns a prediction that follows from his system, which is that in

(3) Each man gave a woman a fun ride and a tasty dish

there is no reading according to which the judge associated with “fun” and the one associated with “tasty” are different persons. I construct such a scenario in which such a reading seems intuitive. So, the fact that this prediction about (3) is born out in Lasersohn's system doesn't speak in its favor.

The last answer to the Binding Argument I consider is that of François Recanati (2002), consisting in appealing to variadic functions. Variadic functions are functions from predicates to predicates whose effect is that of increasing or decreasing the adicity of the input predicate. What is needed are additive variadic functions, which have a twofold role: on one hand they increase the adicity of the input predicate and on the other hand they provide the value for the newly created argument. For example, in a sentence like “John eats in Paris”, “in Paris” is treated as an additive locational variadic operator; its effect in the example given could be captured as follows:

**[Circ<sub>location: Paris</sub> (eats)]** (John) = eats\_in (John, Paris)

For the problematic case (2), the suggestion is to treat the quantifier phrase “every time John lights a cigarette” as an additive locational variadic operator functioning similar to “in Paris” (with the notable difference that now what is provided as the value for the newly created argument is a range of location and not a specific location):

**[Circ<sub>location: every time John lights a cigarette</sub> (rain)]** = rain\_in (the place at which John lights a cigarette every time)

I propose that, in the same vein, one could define an additive variadic operator **Circ<sub>subject</sub>** (call it *subjectual*) having the same role as the additive locational variadic operator in Recanati's view in order to answer to the Binding Argument for PPTs put forward by Schaffer. Thus, in a sentence like “Avocado is tasty for John”, “for John” will be treated as an additive subjectual variadic operator functioning as follows:

**[Circ<sub>subject: John</sub> (tasty)]** (avocado) = tasty\_for (avocado, John).

The problematic sentence (1) is rendered by construing the quantifier “everyone” as a additive subjectual variadic operator, functioning similar to “for John” (with the same qualification in place as in the previous case); formally,

**[Circ<sub>subject: everyone</sub> (tasty)]** (something) = tasty\_for (something, everyone)

This treatment allows the relativist to escape the conclusion of the Binding Argument for PPTs and to claim that unembedded sentences like “Avocado is tasty” don't have an argument place for judges at the appropriate level of representation.

The last section of the paper comprises some common objections to the variadic functions approach and my replies to them.