



VNIVERSITAT DE VALÈNCIA

Doctorado en Ética y Democracia

Neuroética en perspectiva crítica: ¿es posible una
naturalización de la moral desde las neurociencias?

Doctorando: Andrés David Richart Piqueras

Directora: Adela Cortina Orts, Catedrática de Ética y Filosofía Política,
Universitat de València

A Pepe y Puri, mis padres, por haber estado siempre ahí. A Secundina y Miguel, mis tíos, por su incalculable apoyo. A Cristian, mi hermano, por ser un modelo de inteligencia y pensamiento crítico. A Ana Isabel, mi hermana, por toda la ayuda brindada y por sus valiosos consejos. A Eva, mi mejor amiga, por todas esas conversaciones que han sido una inspiración. A Sonia, compañera de vida, por su comprensión y su cariño incondicional. A Maya, mi perra, por acompañarme fielmente en tantas horas de estudio.

También a mis compañeros. Especialmente a Adela y Jesús, por ser un referente, pero sobre todo por su cálido y cordial afecto. Y a Juan Francisco Lisón, por haber creído en mí desde el principio.

Sin vosotros no habría sido posible.

AGRADECIMIENTOS

Diferentes personas e instituciones me han brindado su colaboración y su apoyo, posibilitando el desarrollo de mi investigación y de mi formación como doctor. Deseo expresar mi agradecimiento a todas ellas, reconociendo el importante papel que han desempeñado.

En primer lugar, quiero agradecer a Adela Cortina la inestimable labor que ha llevado a cabo estos años a cargo del Máster Interuniversitario en Ética y Democracia (UV-UJI) y del programa de Doctorado en Ética y Democracia de la Universidad de Valencia, en los que me he formado. También deseo expresar mi agradecimiento por contar con ella como directora de mi tesis doctoral, lo cual me ha dado la oportunidad de trabajar con una persona de su talla académica. Asimismo, quiero agradecer la cercanía y el afecto que durante estos años me ha brindado cordialmente.

En segundo lugar, deseo agradecer a Jesús Conill su apoyo desinteresado, sus valiosos consejos y su inestimable ayuda. Además de ser un referente en el ámbito académico, he encontrado en él un compañero que me ha brindado su afecto cálida y cercanamente.

En tercer lugar, quiero dar mi agradecimiento al Ministerio de Educación Cultura y Deporte, por financiar mi formación y mi investigación mediante una beca de Formación de Profesorado Universitario (FPU).

En cuarto lugar, quiero expresar mi agradecimiento al Ministerio de Industria, Economía y Competitividad, gracias al cual he podido participar en dos proyectos de investigación: “Racionalidad Práctica en Perspectiva Neuroética”, subproyecto “Juicio Moral, Justicia y Democracia en Perspectiva Neuroética” (referencia: FFI2013-47136-C2-1-P); y “Neuroeducación Moral para una Sociedad Pluralista y Democrática”, subproyecto “Neuroeducación Moral, Democracia Deliberativa y Políticas de Desarrollo Humano Sostenible” (referencia: FFI2016-76753-C2-1-P).

En quinto lugar, quiero expresar mi agradecimiento a la Facultad de Filosofía y Ciencias de la Educación de la Universidad de Valencia, al Departamento de Filosofía y a la Unidad Docente de Filosofía Moral, Política y Social, ubicados en dicha facultad, así

como al programa de Doctorado en Ética y Democracia de la Universidad de Valencia, en el que he realizado mis estudios de doctorado.

En sexto lugar, quiero dar las gracias a los compañeros del Departamento de Filosofía y Sociología de la Universidad Jaume I de Castellón. Su disposición a trabajar conjuntamente ha permitido una colaboración muy fructífera y de un incalculable valor para mi formación y mi investigación.

En séptimo lugar, quiero agradecer al Grupo Interuniversitario de Éticas Aplicadas y Democracia (UV-UJI) la oportunidad de unirme a su labor, formando parte de los proyectos de investigación anteriormente mencionados.

En octavo lugar, quiero expresar mi agradecimiento a las instituciones en las cuales complementé mi formación filosófica con formación científica, concretamente en el ámbito de la primatología y la evolución humana. A Fundación Mona (Gerona), en especial a Miquel Llorente y David Riba, he de agradecer la oportunidad de iniciar mi formación en las disciplinas mencionadas, así como el afecto y la disponibilidad con los que siempre me han recibido. También deseo expresar mi admiración por la labor que realizan en calidad de centro de recuperación de primates y en calidad de centro de investigación. Al Departamento de Psicología del Desarrollo y Comparativa, del Max Planck Institute for Evolutionary Anthropology (Leipzig), he de agradecer la oportunidad de realizar una estancia de investigación de carácter internacional.

Finalmente, querría dar mi agradecimiento a todos los compañeros y profesores que he tenido durante este periodo de formación, tanto de la Universidad de Valencia como de la Universidad Jaume I de Castellón. He tenido la inestimable fortuna de encontrar un equipo de personas que dan lo mejor de sí mismas tanto en lo profesional como en lo personal. Gracias a ellas estos años han sido de un incalculable valor.

ÍNDICE

INTRODUCCIÓN	8
OBJETIVOS	12
METODOLOGÍA	14
INTRODUCTION (english version)	16
OBJECTIVES (english version)	20
METHODOLOGY (english version)	22
PARTE I: UNA APROXIMACIÓN CIENTÍFICA A LO MORAL	24
Capítulo 1. Sobre la naturaleza humana y su relación con la moral	26
Capítulo 2. Lo moral en perspectiva filogenética, ontogenética y funcional	36
2.1. Las condiciones naturales de posibilidad del agente moral: origen y desarrollo de la moralidad desde una perspectiva filogenética, ontogenética y funcional	36
2.2. Profundizando en una comprensión de los orígenes de la moralidad	60
Capítulo 3. Sobre la naturalización del juicio moral	76
3.1. El juicio moral desde la perspectiva científica	77
3.2. Errores filosóficos y reduccionistas en la teoría del proceso dual de Greene	98
PARTE II: NEUROÉTICA, ENTRE LA CIENCIA Y LA FILOSOFÍA MORAL	110
Capítulo 4. Moral como estructura y moral como contenido: consideraciones metodológicas	112
4.1. Estructuras, contenidos y fundamentación de lo moral: el error reduccionista de asimilar bases y fundamentos	112
4.2. Libertad como estructura moral: qué puede decir la neurociencia sobre el libre albedrío	130
Capítulo 5. Interdisciplinariedad y método en la neuroética	146
5.1. Retos metodológicos de la neurofilosofía práctica	146
5.2. Materialismo ilustrado como método en neuroética	165
Capítulo 6. Reduccionismo naturalista: perspectiva crítica sobre la naturalización de la moral	190

PARTE III: HACIA UNA METODOLOGÍA CRÍTICA E INTERDISCIPLINAR	210
.....	
Capítulo 7. El problema de la fundamentación y el origen de lo normativo	212
Capítulo 8. Un marco interdisciplinar desde la perspectiva dialógica	236
8.1 Ciencia, ética y el <i>a priori</i> de la comunidad de comunicación	236
8.2 Perspectivas metodológicas y epistémicas en clave dialógica	255
Capítulo 9. De la <i>ethica cordis</i> a una neuroética de la razón cordial	282
9.1. Consideraciones críticas desde la <i>ethica cordis</i>	282
9.2. Neuroética de la razón cordial: criterios metodológicos para una neuroética crítica e interdisciplinar	301
CONCLUSIONES	316
CONCLUSIONS (english version)	322
BIBLIOGRAFÍA	328

INTRODUCCIÓN

En las últimas décadas se ha producido un auge de las neurociencias. Los nuevos avances tecnológicos han facilitado un estudio del funcionamiento cerebral que ofrece la posibilidad de conocer mejor la estructura y el funcionamiento de este órgano. En consecuencia, una multiplicidad de disciplinas ha proliferado bajo la categoría *neuro*. Las posibilidades que se despliegan desde el prisma de las neurociencias son de una variedad enorme y de un alcance aún mayor. Las investigaciones neurocientíficas llevadas a cabo en los últimos años posiblemente no sean más que un prelude de los revolucionarios avances que acontecerán en las próximas décadas.

Entre esas nuevas disciplinas de corte neurocientífico se encuentra la *neuroética*. Definida en términos generales como la investigación ética relacionada con cuestiones neurocientíficas. No obstante, esta definición es demasiado general, cabe hacer distinciones más concretas. Podemos distinguir por un lado la *ética de las neurociencias* o *neuroética aplicada*, que reflexiona en torno a las cuestiones morales implicadas en el desarrollo de las neurociencias, tanto a nivel investigador como a nivel de aplicación. Puede incluirse esta vertiente de la neuroética dentro de las competencias de la bioética. Por otro lado, se encuentran las *neurociencias de la ética*, que tratan de responder algunas de las preguntas fundamentales de la ética mediante el estudio del cerebro. Esta rama es a la que me interesa atender en mi investigación.

Este último frente de la neuroética ofrece una perspectiva muy prolífica para la filosofía moral. Se ocupa de cuestiones tales como qué impacto tiene la investigación neurocientífica en nuestra comprensión de lo que es la ética, cuál es la base neuronal de la agencia moral y qué pueden ofrecer las ciencias del cerebro a la ética en lo que concierne a su fundamentación. Algunas de las preguntas centrales que se han formulado a este respecto son: ¿existen códigos o marcos morales grabados en el cerebro? Y si los hay, ¿son determinantes o solo condicionan la conducta humana? ¿Existe la libertad —entendida como libre albedrío— o el funcionamiento de nuestro cerebro es determinista? ¿Tiene sentido hablar de conceptos como el de autonomía o el de responsabilidad atendiendo a ello? Y, ¿qué consecuencias se derivan de la respuesta a estas preguntas para la educación y la política, así como para el derecho?

Se produce así, en la neuroética, un encuentro entre la ciencia y la filosofía moral. No es la primera vez que esta convergencia se da en la historia, de hecho ciencia y filosofía nacieron siendo una sola cosa, y se han reencontrado numerosas veces. El encuentro, o reencuentro, entre ambas disciplinas viene esta vez propiciado de mano de las neurociencias. Surgen así algunas cuestiones en torno a la relación que en la neuroética mantienen estos dos campos del saber. El estudio de la moral en las neurociencias viene inevitablemente orientado desde el paradigma naturalista propio de las ciencias empíricas. Por su parte la filosofía moral opera mediante la reflexión filosófica, que no se ajusta necesariamente a ese paradigma, aunque atienda en su reflexión a algunas aportaciones de la ciencia. Puesto que la ciencia se adentra en nuestros días en el estudio de lo moral, cabe atender al modo en que ciencia y filosofía moral han de operar cuando comparten el mismo campo de estudio. Es especialmente relevante a este respecto atender al método con el que operan cada una de estas disciplinas.

Las neurociencias pretenden alcanzar un conocimiento objetivo respecto a su objeto de estudio, a través de un método empírico y de un paradigma naturalista. El estudio del cerebro, en tanto que objeto físico y biológico, puede ser abordado de forma pertinente desde las ciencias empíricas; asimismo, las conductas morales también pueden ser estudiadas desde las ciencias empíricas en tanto que cuentan con el cerebro como soporte físico y biológico. Ahora bien, en la medida en que las neurociencias operen en interrelación con la filosofía moral, la cual no trata —o no solo— de lo *que es*, sino de lo *que deber ser*, la cuestión se vuelve más compleja. Esta interdisciplinariedad abre un abanico de posibilidades, pero también de problemas y de cuestiones. ¿Cómo una ciencia empírica que estudia el *es* ha de operar cuando trabaja codo a codo con un campo del saber que reflexiona en torno al *deber ser*? ¿Qué tiene que ofrecer la neurociencia a la filosofía moral? ¿Puede resolver el estudio del cerebro algunos de los problemas fundamentales de la ética? ¿Es capaz la neurociencia de resolver esos problemas por sí misma, prescindiendo de la ética? En caso de que ambas precisen una labor conjunta, ¿qué competencias y qué método corresponden a cada una, y cómo compatibilizarlos? ¿Qué nuevas posibilidades encuentra la ética cuando ejerce su labor en relación con las neurociencias? Estos son algunos de los puntos clave que deben ser aclarados para que la neuroética se desarrolle y se ejerza adecuadamente.

Las neurociencias ofrecen datos sobre cómo el cerebro influye en las conductas morales. Realizan así una descripción de la base neuronal sobre la que se asientan tales conductas, pudiendo atender a su origen evolutivo, a su desarrollo psicológico, a sus funciones y a su estructura, entre otras cosas. Dado que las neurociencias son una rama del conocimiento interdisciplinar, tienen cabida en ellas disciplinas tales como la biología, la psicología, la sociobiología, la etología, y un largo etcétera. A lo largo de la investigación tendré ocasión de atender a estas y otras disciplinas en tanto que suponen un aporte para el estudio neurocientífico de la moral.

Ahora bien, ¿qué significan estos datos para la ética? Los datos descriptivos que las neurociencias ofrecen, ¿qué valor y qué implicaciones tienen para la reflexión ética a la hora de prescribir y justificar normas? ¿Pueden deducirse unas normas o puede abordarse el problema de la fundamentación ética atendiendo a cómo de hecho es la conducta humana a nivel cerebral? Algunos autores intentan plantear una ética fundamental que atienda al cerebro, creen posible hallar unos códigos morales en la estructura neuronal de nuestra especie, que servirían como fundamento para la ética. Otros, con una pretensión más modesta, solo pretenden explicar las bases de la conducta moral humana en tanto que esta tiene el cerebro como soporte físico, pero no llegan a defender que una fundamentación ética sea posible desde una perspectiva meramente empírica y naturalista.

Las neurociencias y la filosofía moral han establecido una simbiosis en la neuroética. Ahora bien, ambas operan en un paradigma diferente y con un método distinto, aquella presenta una perspectiva propia de las ciencias empíricas, esta opera bajo la reflexión filosófica. Hacer compatibles estas dos perspectivas no es fácil. El estudio de la conducta moral humana desde las ciencias supone una naturalización de la moral. En la medida en que la conducta moral se asienta sobre estructuras físicas y biológicas esta naturalización tiene sentido y es legítima. Pero las pretensiones naturalistas de algunos neuroeticistas pueden ser sometidas a crítica, en tanto que las competencias que adjudican a la neurociencia en cuestiones de ética son, cuanto menos, revisables. El auge de las neurociencias viene acompañado en ocasiones de un exceso de optimismo en lo que concierne a sus posibilidades. Si bien las neurociencias tienen mucho que ofrecer a la comprensión del comportamiento moral, se debe ser cauto ante propuestas que abogan

por afrontar los problemas fundamentales de la ética exclusivamente desde el prisma científico.

La investigación que propongo persigue dilucidar cuestiones que precisan ser aclaradas de forma acuciante, la naturalización de la moral en las neurociencias conlleva una serie de problemas cardinales que han de ser formulados de forma explícita, siendo abordados pertinentemente, de lo contrario la labor de la neuroética se llevará a cabo en un marco poco definido en lo que concierne a las competencias y al método que neurociencias y filosofía moral han de asumir cuando ejerzan su labor conjuntamente.

La propuesta metodológica resultante, que será el corolario de la investigación, se articula como una neuroética crítica, concretamente bajo el rótulo de *neuroética de la razón cordial*. Esta propuesta recoge las aportaciones de la ya consolidada Escuela de Valencia, sumándose a ella. Nuestra propuesta de una neuroética de la razón cordial permitirá el ejercicio interdisciplinar de una neuroética crítica en clave cordial, reconociendo la necesidad de atender no solo a los aspectos racionales de la moralidad en la fundamentación de normas y principios morales, sino también a los aspectos emocionales y estimativos, para lo cual las neurociencias ofrecen grandes aportaciones, que son en todo caso necesarias, pero no suficientes, para resolver los problemas fundamentales de la ética.

OBJETIVOS

La hipótesis de la que parte la investigación es que las cuestiones fundamentales de la ética no pueden ser abordadas única y exclusivamente desde el paradigma naturalista. Por ello se propone indagar qué tareas corresponden a las neurociencias y cuáles a la filosofía a la hora de abordar aquellas cuestiones que conciernen a lo moral. Nuestro objetivo será confeccionar una propuesta metodológica para una neuroética crítica, que será articulada, además, en clave cordial. Los objetivos que se pretenden alcanzar a través de la presente investigación son los siguientes:

Objetivo principal: confeccionar una propuesta metodológica para la neuroética que permita resolver los principales problemas asociados al ejercicio de esta, especialmente los relativos a la naturalización de la moral y al reduccionismo naturalista.

Objetivos específicos:

- Objetivo 1: analizar si es posible una naturalización de la moral desde las neurociencias, y de serlo, en qué términos ha de producirse para no incurrir en reduccionismo.
- Objetivo 2: especificar los aspectos generales de la ciencia y de la filosofía moral, su relación, sus competencias, su objeto de estudio y el modo en que ambas deben compatibilizarse en el ejercicio de la neuroética.
- Objetivo 3: analizar las aportaciones más relevantes en el campo de la neuroética para cribar qué aspectos resultan metodológicamente adecuados y cuáles metodológicamente más problemáticos en el ejercicio de esta.
- Objetivo 4: ofrecer una descripción y una explicación de los aspectos naturales de la moralidad desde la perspectiva científica y neurocientífica.
- Objetivo 5: ofrecer criterios para una fundamentación ética de contenidos morales desde la perspectiva de las éticas dialógicas.

METODOLOGÍA

La investigación que llevaremos a cabo es de corte filosófico, por ello la metodología mediante la cual vamos a proceder estará basada en el análisis crítico de las principales aportaciones relevantes para el estudio, incluyendo no solo las propias del campo de la neuroética, sino también todas aquellas que puedan resultar de interés para las diversas cuestiones que hayan de ser abordadas, como, por ejemplo, el problema de la fundamentación o el del reduccionismo naturalista, que ocuparán un lugar central en nuestra investigación.

En primer lugar, pretendo aclarar qué se entiende por naturalismo en el marco de la neuroética. Cabe aquí atender a las propuestas de diferentes neuroeticistas y filósofos para analizar qué concepto de naturalismo manejan. Cabe atender también al concepto de naturaleza humana. En sintonía con cómo se haya definido esa naturalización, ese naturalismo y esa naturaleza humana, pretendo esclarecer qué posibilidades ofrecen las neurociencias en torno a las cuestiones de las que se ocupa la ética, especialmente en lo que concierne a la fundamentación. Para ello haré un análisis de las principales aportaciones que el estudio del cerebro ofrece respecto a las conductas morales humanas. En relación a ello cabe atender a diferentes campos en los que las neurociencias son especialmente fértiles, siendo centrales cuestiones como la filogénesis de la moral, su desarrollo ontogenético, el estudio del funcionamiento cerebral, el juicio moral, el libre albedrío, etc. El primero de ellos trata de comprender cómo se gestaron evolutivamente los comportamientos sociales que, más tarde, desembocaron en las conductas morales de nuestra especie. El segundo aborda la cuestión de cómo se produce el desarrollo de lo moral en la ontogénesis, entendida como desarrollo psicológico en relación con el medio durante la vida. Esa evolución y desarrollo comportamental se produce en correlación a cambios cerebrales, y su comprensión es uno de los grandes logros a los que aspira la neuroética. El tercero de esos campos se ocupa de investigar el funcionamiento del cerebro, su estructura, su desarrollo epigenético y las relaciones que estos guardan con la conducta moral. El cuarto concierne al estudio neurocientífico del juicio moral. El quinto refiere a la cuestión de si la conducta humana está determinada o si existe el libre albedrío, siendo este una de las condiciones de posibilidad de la moralidad.

Atenderé, además, a diferentes propuestas de neuroética con el objeto de ver qué suponen para la filosofía moral las aportaciones de las neurociencias, atendiendo al método y a las competencias correspondientes a cada una de ellas. El estudio de propuestas concretas sobre la metodología propia de la neuroética permitirá analizar qué problemas y virtudes presentan algunos de los enfoques metodológicos más representativos en la disciplina. Con ello se pretende ofrecer respuestas a los interrogantes que suscitan esta investigación.

Tras ello, analizaremos el papel de la filosofía moral en el ejercicio de la neuroética. En este punto la investigación se inscribe en una corriente muy concreta, la de la ética del discurso, heredera del pensamiento kantiano y de la Teoría Crítica. Desde esta corriente analizaremos la relación existente entre ciencia, filosofía moral y sus respectivas metodologías. Desembocaremos en una propuesta metodológica que supere las dificultades a las que se enfrenta la neuroética, esta propuesta será marcadamente interdisciplinar, reconociendo a la ciencia y a la filosofía moral las competencias que le son propias, permitiendo configurar una neuroética que sea fiel a su nombre, siendo una unión de las neurociencias y de la ética, y no meramente una neurociencia de lo moral, desprovista del componente filosófico, del cual no se puede prescindir sin caer en reduccionismo naturalista.

INTRODUCTION (english version)

In the last decades there has been a boom in neurosciences. The new technological advances have facilitated the study of the cerebral functioning that offers the possibility to know the structure and the functioning of this organ better. In consequence, a multiplicity of disciplines has proliferated under the ‘neuro’ category. The possibilities that are deployed from the prism of the neurosciences are of a huge variety and of an even greater scope. The neuroscientific research carried out in recent years may be no more than a prelude to the revolutionary advances that will take place in the coming decades.

Among these new neuroscientific disciplines is *neuroethics*. Defined in general terms as the ethical research related to neuroscientific issues. However, this definition is too general, more specific distinctions should be made. On one hand, we can distinguish *the ethics of neuroscience* or *applied neuroethics*, which reflects on the moral issues involved in the development of neurosciences, both at the research level and at the implementation level. This aspect of neuroethics can be included within the competencies of bioethics. On the other hand, there are the *neurosciences of ethics*, which try to answer some of the fundamental questions of ethics through the study of the brain. This field is the one that I am focusing in my research.

This last aspect of neuroethics offers a very prolific perspective for moral philosophy. It deals with questions such as what impact does neuroscientific research have on our understanding of what ethics is, which is the neural basis of morality, and what can brain sciences offer ethics regarding its grounds. Some of the main questions that have been asked on this matter are: are there moral codes or frames engraved in the brain? And if so, are they determining or only conditioning of the human behaviour? Is there freedom—understood as free will—or the functioning of our brain is deterministic? Does it make sense to talk about concepts such as autonomy or responsibility in response to it? And, what consequences stem from the answer to these questions for education and politics, as well as for law?

Therefore, there is an encounter between science and moral philosophy in neuroethics. It is not the first time that this convergence occurs in history, actually science and philosophy were born as a single thing and they have meet numerous times. The encounter, or reencounter, between both disciplines comes propitiated by the

neurosciences this time. Therefore, some questions regarding the relation that these two disciplines have in neuroethics arise. The study of morality in the neurosciences is inevitably oriented from the naturalistic paradigm of the empirical sciences. For its part, moral philosophy operates through philosophical reflection, which does not necessarily conform to that paradigm, even if it pays attention to some science contributions in its reflection. Since nowadays science goes into the study of morality, we must consider the way in which science and moral philosophy have to operate when they share the same field of study. It is especially relevant in this regard to pay attention to the method with which each one of these disciplines operate.

The neurosciences aim to achieve objective knowledge about their object of study, through an empirical method and a naturalist paradigm. The study of the brain, as a physical and biological object, can be approached in a relevant way from the empirical sciences; likewise, moral behaviours can also be studied from the empirical sciences as they count on the brain as a physical and biological support. However, to the extent that neurosciences operate in interrelation with moral philosophy, which does not deal, or not only, with *what it is* but with *what it should be*, the question becomes more complex. This interdisciplinarity opens a range of possibilities, but also of problems and issues. How does an empiric science that studies *what it is* operate when it works hand in hand with a field of knowledge that reflects on *what it should be*? What does neuroscience have to offer to moral philosophy? Can the study of the brain solve some of the fundamental problems of ethics? In case both need a joint work, what competences and what method correspond to each one and how to make them compatible? What new possibilities does ethics encounter when it works with neurosciences? These are some of the key points that need to be clarified so that neuroethics can be developed and practiced properly.

Neuroethics offer data about how brain influences moral behaviours. They make a description of the neuronal base on which such behaviours are based, being able to take into consideration their evolutionary origin, their psychological development, their functions and their structure, among other things. As neurosciences are a field of interdisciplinary knowledge, disciplines such as biology, psychology, sociobiology, ethology and so on have a place in them. Throughout the research, I will have occasion to pay attention to these and other disciplines as far as they contribute to the neuroscientific study of morality.

However, what do this data mean to ethics? What value and what implications do the descriptive data that neuroethics offers have to ethical reflection when prescribing and justifying norms? Can some norms be deduced or can the problem of ethical foundation be addressed in light of how human behaviour actually is at the brain level? Some authors try to pose a fundamental ethic that pays attention to the brain, believing in the possibility of finding moral codes in the neural structure of our species, which would serve as a foundation for ethics. Others, with a more modest ambition, just try to explain the bases of the human moral behaviour as much that this has the brain as a physical support, but they do not defend that an ethical foundation is possible from a merely empirical and naturalistic perspective.

Neuroscience and moral philosophy have established a symbiosis in neuroethics. However, they work in different paradigms and with different methods; neuroscience has a typical empiric sciences perspective whether moral philosophy works under philosophical reflection. Making both compatible is not easy. The study of human moral behaviour from the sciences implies a naturalization of morality. To the extent that moral behaviour is based on physical and biological structures this naturalization makes sense and is legitimate. But the naturalistic pretensions of some neuroethicists can be subjected to criticism, while the competences that they attribute to neuroscience in matters of ethics are, at least, revisable. The rise of neuroethics sometimes comes with an excess of optimism regarding its possibilities. Although neurosciences have a lot to offer to the understanding of moral behaviour, caution should be exercised before proposals that advocate addressing the fundamental problems of ethics exclusively from the scientific prism.

The research that I propose aims to elucidate issues that need to be clarified urgently, the naturalization of morality in the neurosciences involves a series of cardinal problems that have to be formulated explicitly, being addressed pertinently; otherwise, the work of neuroethics will be carried out in a frame that is not well defined in what concerns the competences and the method that neurosciences and moral philosophy have to assume when they work together.

The resulting methodologic proposal, that will be the corollary of the research, is articulated as a critical neuroethics, specifically under the label of *neuroethics of cordial reason*. This proposal collects the contributions of the already consolidated School of

Valencia, joining it. Our proposal of a neuroethics of cordial reason will allow the interdisciplinary practice of a critical neuroethics in a cordial key, recognising the need to pay attention not only to the rational aspects of morality in the foundation of norms and moral principles, but also, to the emotional and estimative aspects which are in any case necessary, but not sufficient, to solve the fundamental problems of ethics.

OBJECTIVES (english version)

The hypothesis from which the research is based is that the fundamental issues of ethics can not be approached solely and exclusively from the naturalist paradigm. Therefore, it is proposed to investigate which tasks correspond to the neurosciences and which to philosophy when addressing those issues that concern morality. Our objective will be to compose a methodological proposal for a critical neuroethics, which will also be articulated in a cordial key. The objectives that are intended to be achieved through the present research are the following:

Main objective: to compose a methodological proposal for neuroethics that allows to solve the main problems associated to its practice, especially those related to the naturalization of morality and to the naturalistic reductionism.

Specific Objectives:

- Objective 1: to analyse if it is possible a naturalization of morality from the neurosciences, and if so, in what terms it must occur in order not to incur reductionism.
- Objective 2: to specify the general aspects of science and moral philosophy, their relationship, their competences, their object of study and the way in which both need to be compatible in the practice of neuroethics.
- Objective 3: to analyse the most relevant contributions in the field of neuroethics in order to determine which aspects are methodologically appropriate and which methodologically most problematic in the practice of it.
- Objective 4: to offer a description and an explanation of the natural aspects of morality from a scientific and neuroscientific perspective.
- Objective 5: to offer criteria for an ethical foundation of moral content from the perspective of dialogical ethics.

METHODOLOGY (english version)

The research that we will carry out is philosophical. Therefore, the methodology through which we will proceed will be based on the critical analysis of the main relevant contributions to the study including not only those of the field of neuroethics, but also all those that may be of interest for the various issues that have to be addressed, such as, for example, the problem of the foundation or the naturalistic reductionism, which will occupy a central place in our research.

First of all, I intended to clarify what is understood by naturalism in the framework of neuroscience. It is appropriate here to pay attention to the proposals of different neuroethicists and philosophers to analyse what concept of naturalism they use. It is also appropriate to address the concept of human nature. In line with how this naturalization, this naturalism and that human nature has been defined, I intend to clarify what possibilities do neurosciences offer around the issues dealt with by ethics, especially as regards the foundation. For this, I will make an analysis of the main contributions that the study of the brain offers regarding human moral behaviours. In relation to this, we must pay attention to different fields in which neurosciences are especially fertile, being questions such as the phylogenesis of morality, its ontogenetic development, the study of brain functioning, moral judgment, free will, etc. The first of them tries to understand how did the social behaviours that later led to the moral behaviours of our species arose and evolved. The second one addresses the question of how the development of morality occurs in ontogenesis, understood as psychological development in relation to the environment during life. This evolution and this behavioural development occur in correlation to brain changes, and its understanding is one of the great achievements to which neuroethics aspires. The third one of these fields investigates the functioning of the brain, its structure, its epigenetic development and the relationships that these have with moral behaviour. The fourth one concerns the neuroscientific study of moral judgment. The fifth one refers to the question of whether the human behaviour is determined or if free will exists, being free will one of the conditions of possibility of morality.

I will also study different proposals of neuroethics in order to see what the contributions of the neurosciences represent for moral philosophy, taking into account the method and the competences corresponding to each one of them. The study of specific

proposals about the methodology of neuroethics will allow to analyse what problems and virtues some of the most representative methodological approaches in the discipline present. This is intended to provide answers to the questions raised by this research.

After this, we will analyse the role of moral philosophy in the practice of neuroethics. At this point, the research is part of a very specific current, that of discourse ethics, heir to the Kantian thought and the Critical Theory. From this current, we will analyse the existing relationship between science, moral philosophy and their respective methodologies. We will end in a methodological proposal that overcomes the difficulties faced by neuroethics, this proposal will be markedly interdisciplinary, recognising its own competences to science and moral philosophy, allowing us to configure a neuroethics that is true to its name, being a union of neurosciences and ethics, and not merely a neuroscience of the morality, devoid of the philosophical component, which can not be disregarded without falling into naturalistic reductionism.

PARTE I: UNA APROXIMACIÓN CIENTÍFICA A LO MORAL

[...] todo animal, cualquiera que sea su naturaleza, si está dotado de instintos sociales bien definidos, incluyendo entre éstos las afecciones paternas y filiales, inevitablemente llegaría a la adquisición del sentido moral o de la conciencia cuando sus facultades intelectuales llegasen o se aproximasen al desarrollo a que aquellas han llegado en el hombre.

Charles Darwin, *El origen del hombre y la selección en relación al sexo*

Dado que las neurociencias así como otras disciplinas científicas muestran a día de hoy un creciente interés por la cuestión de lo moral se hace preciso aclarar algunos interrogantes, se han convertido en preguntas centrales tanto para la filosofía como para la ciencia algunas cuestiones como qué entendemos por naturaleza y en concreto qué entendemos por naturaleza humana, en qué sentido esta puede ser relevante para el estudio de la moral, qué puede la ciencia descubrir sobre la moral y qué implicaciones tienen esos descubrimientos para la ética.

El punto de partida adecuado para la investigación que aquí iniciamos debe tomar como eje de la reflexión la cuestión de la naturaleza moral humana. Algunos de los objetivos que pretendemos resolver en este punto de la investigación son: hacer una aclaración terminológica sobre el concepto de naturaleza que vamos a manejar, que será aquel vinculado a las ciencias empíricas en tanto que naturalistas; aclarar en el mismo sentido el concepto de naturaleza humana; analizar los diferentes modos en que los resultados de las investigaciones científicas en torno a esa naturaleza humana se relacionan con la ética; seleccionar cuáles de esos modos son relevantes para nuestra investigación y por qué; analizar qué implicaciones tienen para la ética los resultados de la investigación científica en torno a la moral; y exponer el modo en que las ciencias empíricas han abordado el estudio de lo moral, siendo esta cuestión ampliada en capítulos posteriores.

Capítulo 1. Sobre la naturaleza humana y su relación con la moral

El ser humano en tanto que ser natural es susceptible de ser tomado como objeto de estudio de la ciencia. En qué consiste la agencia moral, entendida como facultad que permite la moralidad, es algo que puede ser explicado por la ciencia en la medida en que es la constitución natural de nuestra especie la que favorece su aparición. Una de las tesis que defenderé más adelante en la investigación es que las facultades del ser humano, naturalmente constituidas, permiten la posibilidad de un espacio de reflexión sobre objetos que no están en el mundo, pero que tratan de orientar la acción en este, atendiendo a ciertas ideas y conceptos a los que llamamos morales. La comprensión de esta brecha entre las condiciones de posibilidad de la agencia moral y el espacio de reflexión ética que acontece en el agente moral permitirá dirimir qué competencias escapan a la ciencia en cuestiones de filosofía moral. Para un adecuado estudio de la cuestión se esclarecerá previamente a lo largo de este capítulo en qué sentido emplearemos los conceptos de naturalismo, naturaleza humana y naturaleza moral humana, así como los modos en que la naturaleza humana se relaciona o resulta relevante para lo moral.

¿Naturalizar la ética?

En las últimas décadas hemos experimentado un desarrollo sin precedentes en el estudio científico del cerebro humano, a pesar de que ya se habían dado algunos pasos en este campo. El rótulo *neuro* inunda los ámbitos de investigación y divulgación no solo científica, sino de todas las áreas del saber (Siurana, 2017: 109, 199 y 201). La interdisciplinariedad de las neurociencias ha convulsionado ramas del conocimiento tales como el derecho, la política o la ética, entre muchas otras, así como diferentes ámbitos de la sociedad, como por ejemplo el de la empresa (González Esteban, 2016). La filosofía y la ciencia, que estaban unidas en su origen, han tenido algunos reencuentros a lo largo de la historia, esta vez la concurrencia entre ambas viene propiciada desde las neurociencias (Cortina, 2012a; Salles y Evers, 2017). Los estudios neurocientíficos reclaman por derecho propio la investigación de aquellas cuestiones que puedan ser abordadas mediante su método, a saber, aquellas cuestiones cuya explicación es accesible mediante un adecuado estudio de los hechos, a través de la experiencia, formulando después hipótesis que puedan dar cuenta de tales hechos y que deberán resistir la crítica de la

comunidad científica (Habermas, 2016; Cortina, 2007). Al complejísimo problema de interpretar y explicar los hechos de forma acertada se suma otro que es, si cabe, más complejo aún. Mientras muchos neurocientíficos se aventuran en buscar respuestas, se va haciendo cada vez más patente que antes de poder dar buenas respuestas hay que distinguir muy bien qué tipo de preguntas pueden ser en efecto respondidas desde el paradigma naturalista de la ciencia. Si bien la ciencia es una de las mejores bazas con las que cuenta el ser humano en su búsqueda del conocimiento, eso no significa que pueda responderlo todo ni que sea infalible (Buchanan, 2009; Apel, 1985; Habermas, 2006; Aranguren, 1994; Gould, 1999; Hume 1988, 1953; Moore, 1983; Salles, 2013, 2016, 2018; Jácomo, 2011, 2013a, 2013b, 2014; Evers, 2010; Cartwright, 2010; Teehan y diCarlo, 2004; Cortina, 2012a; García-Marzá, 2016; Conill, 2012).

En lo que concierne a la filosofía moral, que es uno de los temas que nos ocupa, las primeras preguntas que nos asaltan tienen que ver con cuestiones clásicas de la ética, siendo entre ellas de especial relevancia la pregunta por la fundamentación. ¿Podrá la neurociencia dar una fundamentación, en sentido de justificación normativa, de unos códigos éticos que permitan regular moralmente la conducta humana?, ¿están esos códigos grabados en el cerebro humano o solo encontraremos en este una serie de predisposiciones conductuales configuradas por nuestra psicobiología? Si la ciencia se propone adentrarse en cuestiones que tradicionalmente habían sido abordadas por la filosofía moral, hemos de plantearnos qué es exactamente la ciencia, qué puede hacer, cómo lo hace, y qué queda fuera de sus posibilidades.

Tendremos ocasión de abordar de forma amplia la cuestión de cómo comprendemos la ciencia en la Parte III de la presente investigación, por ahora atenderemos a una definición sumaria de la misma, que retomaremos y ampliaremos en su debido momento. La ciencia puede definirse como un conjunto de conocimientos sistemáticamente estructurados. Estos conocimientos se componen de enunciados que tienen una pretensión de verdad respecto a hechos que se ubican en el mundo. El quehacer científico se basa en la observación y en la experimentación. Todo conocimiento científico ha de poder contrastarse con los hechos mediante la experiencia, es un conocimiento empírico. No obstante, los hechos no hablan por sí solos, los científicos han de interpretarlos, formular hipótesis y construir teorías, han de dar forma a los datos (Gadamer, 1991; Conill, 2006, 2015; Apel, 1985; Habermas, 2006; Cortina, 2007,

2012a). Aquí se introducen varios elementos relevantes. Uno de ellos es el de la interpretación, la ciencia no es un reflejo o un acceso directo al mundo natural. Se introduce además el razonamiento como uno de los elementos que componen su método. Las teorías y las hipótesis serán vigentes siempre que resistan la crítica, de este modo la pretensión de objetividad cristaliza en la intersubjetividad de la comunidad científica, que mediante consenso dialógico acuerda qué podemos considerar como conocimiento, siendo este vigente, pero revisable y susceptible de ser desestimado de forma futura (Habermas, 2006; Apel, 1985; Cortina, 2007).

La naturaleza, entendida como todo aquello que compone el mundo material, es el objeto de estudio de la ciencia. Ésta opera desde un paradigma naturalista que toma la naturaleza como principio explicativo de todas las cosas, si bien no descarta la validez de cualquier conocimiento analítico y formal, llegando incluso a servirse de este tipo de conocimiento como herramienta auxiliar (Apel, 1985; Cortina, 2007).

Respecto a lo sobrenatural, el naturalismo científico no necesariamente lo niega, sin embargo las posturas más concesivas que la ciencia puede adoptar no pasan de declaraciones de agnosticismo respecto a realidades sobrenaturales. Ahora bien, no debemos quedarnos en los polos de lo natural opuesto a lo sobrenatural. La filosofía moral se pregunta por el origen de la obligación moral, que se presenta al agente moral como incondicional (Kant, 1992, 2000; Cortina; García-Marzá; Apel; Habermas). La ciencia puede dar cuenta de la base biológica y cerebral que subyace al sentimiento humano de obligación, pero sin duda la obligación en sí es otra cosa, tiene un sentido que escapa a la mera explicación naturalista y empírica. Es aquí donde el trascendentalismo —bien kantiano (Kant) o bien en las contemporáneas formas de ética dialógica (Apel; Habermas; Cortina; García-Marzá)— cobra especial relevancia para comprender qué es la moral, pudiendo dar cuenta de cuestiones para las cuales el naturalismo empirista es insuficiente, sin apelar por ello a entidades sobrenaturales. La ciencia, pues, no puede asumir todas las tareas de investigación y resolución de problemas en lo que respecta a la ética.

El ser humano es un ser natural, se ubica en el mundo material y forma parte de él, sus sentimientos, sus pensamientos, lo que conoce, etc., son elementos que se alojan en una estructura biológica material, en su cuerpo, en sus genes, en su cerebro. Puesto que el ser humano es un ser natural es susceptible de ser tomado como objeto de estudio desde las ciencias empíricas. Pero ¿qué hay de esas otras realidades que, como la

obligación moral, no pueden ser entendidas meramente desde el naturalismo empírico? ¿Es el ser humano un ser natural en el que se dan las condiciones para que emerjan elementos con base material pero que no son reductibles a lo material? El problema no podría ser más central. Conocer la naturaleza del ser humano debería darnos claves sobre qué compete a la ciencia en el estudio de este y qué no, y en lo que respecta a su naturaleza moral, debería facilitar la comprensión de qué cuestiones de la ética pueden ser naturalizadas y cuáles no.

¿Cómo entendemos la naturaleza humana?

Históricamente se han ofrecido distintas versiones de aquello que podemos entender por naturaleza humana, Aristóteles, Hobbes y Hume representan solo algunas de las figuras que trataron de explicar en qué consiste la naturaleza de nuestra especie, y todos lo hicieron de diferente modo. No me ocuparé aquí de recapitular estas u otras propuestas en torno a la cuestión, creo que es mucho más productivo tratar de acotar qué concepto de naturaleza humana es aquel que nos interesa a la hora de analizar la relación entre neurociencias y filosofía moral.

El concepto de naturaleza humana ha recobrado importancia en las últimas décadas a propósito del debate sobre la mejora humana. Aprovechando la coyuntura de este debate, pero sin adentrarnos en él, podemos tomar algunas de las definiciones más actualizadas del concepto. Atenderemos para ello a diferentes autores, Allen Buchanan, Norman Daniels y T. M. Murphy. El trabajo de Buchanan y Daniels en lo que refiere al concepto de naturaleza humana y al problema de la biomejora es reconocido por autores emblemáticos en el debate, Habermas, sin ir más lejos, reconoce el mérito de las investigaciones de estos autores en una nota al pie de página en *El futuro de la naturaleza humana* (Habermas, 2002: 37).

En un reciente artículo Buchanan ofrece algunas definiciones de naturaleza humana que resultan interesantes para nuestra investigación, a saber: *philosophical*, *folk*, *evolutionary*, y *religious*. Entre ellas nos interesa atender a las de corte filosófico y

evolucionista¹, puesto que plantean el concepto en el sentido que nos interesa. Veamos la definición filosófica del concepto:

[...] la concepción filosófica dominante de la naturaleza humana consiste en un conjunto de características que son comunes a todos los humanos y que distingue a los humanos de otras clases de seres. Si uno añade la asunción de que hay clases naturales, entonces esas características son pensadas como esenciales, en lugar de simplemente contingentes. (Buchanan, 2009: 142)

Atendamos ahora a la definición evolucionista:

[...] Algunos biólogos evolutivos y neurocientíficos contemporáneos dirían que si el concepto de naturaleza humana tiene algún valor es como ese conjunto de características innatas que la mayoría de humanos ahora tienen como resultado de su desarrollo evolutivo común. Común tanto a la concepción popular como a la evolutiva es la idea de que si algo es parte de la naturaleza humana es reacio a la alteración por aculturación, educación o indoctrinación, y estable a través de una amplia variedad de entornos. (Buchanan, 2009: 142)

Estas dos definiciones de naturaleza humana, que no son excluyentes entre sí y que se complementan, representan los rasgos básicos del concepto. Según Buchanan, una perspectiva filosófica y evolucionista de la naturaleza humana definiría esta del siguiente modo: conjunto de características comunes a todos los humanos, y presente solo en ellos, lo cual permite diferenciarles de otras especies, ese conjunto de características ha sido adquirido como resultado de un desarrollo evolutivo común y está presente en todos los miembros de nuestra especie independientemente de las contingencias culturales y educativas que les diferencian epigenética y ontogenéticamente.

Timothy F. Murphy —que comenta algunos puntos de la propuesta de Norman Daniels— confirma algunos de los aspectos de esta definición en otro artículo de la misma revista: “la naturaleza de un organismo refiere a sus rasgos característicos, que nunca exhiben variación por razones genéticas o ambientales” (Murphy, 2012: 191). Ambos autores manejan un concepto descriptivo de naturaleza humana, propio de una perspectiva naturalista. Murphy señala en ese mismo texto la distinción existente entre las perspectivas descriptivas y las metafísicas respecto al concepto (Murphy, 2012: 195).

¹ Respecto a las otras dos definiciones, *folk* y *religious*, no atenderemos a ellas porque la primera es una noción vaga de naturaleza humana y la segunda incluye una perspectiva que se aleja del propósito que aquí perseguimos.

El desarrollo de la ciencia moderna ha propiciado una serie de avances que nos permite conocer al ser humano en tanto que objeto perteneciente al mundo natural. Es en este contexto donde el concepto de naturaleza humana se va configurando en sentido meramente descriptivo. La perspectiva científica toma al ser humano como un objeto más perteneciente a ese mundo material que constituye su objeto de estudio. Este será el sentido en el que tomaremos el concepto en la presente investigación: cada ser humano es un ser perteneciente al mundo material, en tanto que es así presenta una serie de características que componen su naturaleza y que son susceptibles de ser tomadas como objeto de estudio científico, siendo su conocimiento accesible mediante la experiencia y desde una perspectiva descriptiva y explicativa; dado que la especie humana está compuesta por individuos concretos, la naturaleza humana será una abstracción de aquellas cualidades naturales que presumiblemente sean compartidas en mayor o menor medida por la mayoría de seres humanos adecuadamente constituidos, independientemente de las variables que intervengan en su desarrollo epigenético y ontogenético —culturales, sociales, familiares, educativas, etc.

La naturaleza humana y su relevancia para lo moral

La complejidad del ser humano hace que en tanto que ser natural presente toda una constelación de dimensiones, interrelacionadas entre sí, que componen eso que llamamos su naturaleza. Nuestro objetivo va más allá de comprender a qué nos referimos al hablar de naturaleza humana en el ámbito científico, lo que queremos es atender a un aspecto concreto de esta, a la dimensión moral de la misma. El conocimiento de la naturaleza moral humana, en el sentido descrito en el apartado anterior, parece hoy en día más accesible que nunca gracias a los avances en el estudio del cerebro humano. La neurociencia, que es interdisciplinaria, integra además los logros de la biología, la sociobiología, la etología, la psicología, la genética, etc., componiendo una gama muy rica de elementos que ofrecen la posibilidad de esbozar qué es eso que hace moral a nuestra especie, por qué somos morales, y si todos los aspectos de la ética son reductibles o no a la base material que forma al ser humano. ¿Qué puede decirnos el estudio de la naturaleza humana acerca del ser humano en tanto que ser moral?, ¿qué puesto ocupa la ética en relación con la naturaleza humana?

Retomamos aquí el artículo de Buchanan, en él el autor hace algunas observaciones que establecen un buen punto de partida para la reflexión, señala cinco papeles o roles que la naturaleza humana juega en relación a la ética:

Cinco diferentes roles han de ser distinguidos: (1) NH como condición de la agencia moral, (2) NH como restricción de viabilidad en la moralidad, (3) NH como restricción en lo bueno para los humanos, (4) NH como una fuente de reglas morales sustantivas, y (5) NH como un todo -un complejo juego de características interdependientes- que puede ser seriamente dañado por los esfuerzos en mejorarla. (Buchanan, 2009: 143)

En primer lugar, el autor señala que la naturaleza humana juega un papel en la ética como condición de posibilidad para la agencia moral. La constitución biológica, genética, neuronal, psicológica, etc., de nuestra especie compone las condiciones de posibilidad necesarias para que el ser humano pueda constituirse como agente moral, sentando las bases materiales que desde su biología posibilitan este hecho. Esto quiere decir que sin la dimensión natural del ser humano —o la de algún otro ser con facultades análogas, en caso de que exista o existiese— no cabría posibilidad alguna de que hubiese espacio para la ética. La ética precisa de unas condiciones naturales para poder existir, y el ser humano las aglutina. Ahora bien, esto no significa que todas las cuestiones de la ética puedan abordarse estudiando las condiciones naturales que son necesarias para su existencia, la ética es algo más que la suma de condiciones que la hace posible. Qué condiciones son estas que reúne la naturaleza humana y cómo posibilitan la ética será un tema que retomaremos más adelante para profundizar de lleno en él.

En segundo lugar, la naturaleza humana se pone en relación con la ética como una restricción de viabilidad sobre la moralidad. Los últimos avances en psicología moral y neurociencia están mostrando que los seres humanos estamos fuertemente condicionados por patrones conductuales que imponen ciertos límites a nuestra conducta moral (Cortina, 2012a; Evers, 2010; Damasio; Greene; Churchland; Forbes, 2010; Frith, 2008; Salles, 2017). Existe, por ejemplo, una tendencia a la endogamia moral (Evers, 2010; Tomasello, 2013; Waal, 1997, 2007; Kohlberg, 1992; Cortina, 2012a, 2017; Greene, 2013; Churchland, 2013).

Algunos filósofos morales, incluyendo Hume y practicantes contemporáneos de la ‘neuroética’, enfatizan que una comprensión realista adecuada de la moralidad debe tener en cuenta las limitaciones cognitivas y motivacionales de los seres humanos (en jerga de la neuroética, el ‘cableado duro’ [hard-wiring] que de hecho tenemos como resultado de la evolución). Si, por

ejemplo, en virtud de nuestra estructura biológica evolucionada, tenemos una capacidad limitada para actuar altruistamente hacia los extraños o hacia aquellos que no son miembros de nuestro ‘grupo primario’, entonces una explicación plausible de nuestras obligaciones morales hacia otros debería tener esto en cuenta. (Buchanan, 2009: 143)

Si los resultados de las investigaciones científicas están en lo cierto, si existen por naturaleza unos límites irrebasables en nuestra constitución psicológica y biológica, entonces no tendremos que investigar solamente qué normas han de regir nuestra conducta y nuestras sociedades, sino que además deberemos de hallar el camino de hacer viable su aplicación y cumplimiento dentro de los límites que nuestra naturaleza impone, o bien aplicar técnicas de biomejora moral para mover los hitos que marcan esos límites. Lo que se plantea aquí no es un reduccionismo que pretenda deducir de nuestra naturaleza una serie de normas morales, al contrario, se está advirtiendo de las dificultades de aplicación que presentan algunas normas y principios morales en la medida en que ponerlos en práctica sea algo inviable dada nuestra constitución natural. Este es un interesantísimo campo de reflexión que gracias a los avances en neurociencias promete ser muy prolífico en los próximos años. A la necesidad de buscar los principios morales que regulen nuestra conducta se suma el imperativo de buscar modos viables de aplicación de esas exigencias normativas, que no pueden desatender la naturaleza de nuestra especie en la medida en que esta impone límites. Cuáles son los límites que nuestra neurobiología impone es ahora mismo objeto de estudio de las neurociencias, la plasticidad del cerebro durante el desarrollo epigenético permite una gran flexibilidad, aún están por descubrir y acotar adecuadamente esos límites (Evers, 2010, 2015; Salles, 2017; Jácomo y Serrão, 2013).

Un tercer papel que nuestra naturaleza juega en relación con la ética es como restricción de lo que es bueno para nosotros. Esta perspectiva es recogida hoy en día por propuestas neoaristotélicas como el enfoque de las capacidades (Sen, 1979; Nussbaum, 2006; MacIntyre, 2001).

En una plausible interpretación, Aristóteles sostuvo que la naturaleza de un ser determina su bien, pero no solo en un sentido mínimo: por constituir una restricción en lo que puede contar como una vida buena para esa clase de ser. (Buchanan, 2009: 144)

La naturaleza de un ser tendrá, pues, mucho que ver con qué es aquello que resulte una vida buena para ese ser. En función de ello cabe tomar en cuenta la naturaleza de un

ser a la hora de establecer las pautas que regirán nuestra conducta en la medida en que esta le afecte, dicho de otro modo, qué sea aquello que resulte bueno para ese ser es algo que tendrá mucho que ver con su naturaleza. Ahora bien, hay que evitar toda tentativa de reduccionismo, evitar el error deductivo de considerar que sus atributos naturales son intrínsecamente normativos y que de ellos se pueden extraer directamente una serie de conclusiones normativas (Buchanan, 2009; Apel, 1985; Habermas, 2006; Hume 1988, 1953; Moore, 1983; Cartwright, 2010; Teehan y diCarlo, 2004; Cortina, 2007, 2012a; García-Marzá, 2016; Koorsgard, 2000).

En cuarto lugar, advierte Buchanan, la naturaleza humana puede ser tomada erróneamente como fuente de normas morales sustantivas. El autor se muestra crítico con esta postura, a la que denomina *esencialismo normativo*: “ellos creen que es posible derivar reglas morales sustantivas a partir de la observación de la naturaleza humana” (Buchanan, 2009: 145). Esta pretensión incurre en reduccionismo naturalista, partiendo del hecho de cómo sea la naturaleza humana no puede derivarse un deber ser en sentido moral. A lo largo de la investigación seremos críticos con las pretensiones de esta índole, ya que vician el buen hacer en la neuroética, dado que atribuyen a las ciencias empíricas unas competencias que no le corresponden.

Por último, señala, la naturaleza humana puede ser entendida como una compleja red de elementos interdependientes, esta apreciación resulta relevante para la discusión sobre la mejora humana, puesto que alterar una pequeña parte de nuestra naturaleza afectaría a toda ella, sin embargo no constituye uno de los puntos centrales para nuestra investigación.

Los cinco roles que Buchanan señala suponen una referencia valiosa, pero no recogen todos los posibles papeles que la naturaleza humana puede jugar en relación a la moral. Farah señala otro posible rol, de una importancia esencial. Según indica la autora, precisamente la naturaleza de nuestra especie puede servir como motivo para atribuir un valor moral especial a sus miembros. Si hay motivos para considerar dignos en sentido moral a los seres humanos —o algún otro tipo de seres—, entonces será en tanto que la naturaleza de tales seres reúne las características necesarias para atribuir ese reconocimiento de valor (Farah; Cortina; Kant).

Del mismo modo que tradicionalmente hemos visto a las personas como diferentes de otros objetos debido a su capacidad de agencia moral, también las hemos visto como teniendo un valor moral especial, a diferencia de todos los otros tipos de objetos. Las personas merecen protección contra el daño solo porque son personas. (Farah, 2012: 587)²

Entre los seis papeles mencionados que la naturaleza humana puede jugar en relación con lo moral no todos resultan igual de relevantes para nuestra investigación. En este sentido merecen especial atención los roles relativos a la naturaleza humana como condición de posibilidad de la moral, como elemento para la fundamentación de principios y normas, que tomaremos en perspectiva crítica, y como principio para el reconocimiento de valor moral o dignidad.

La cuestión que pretendemos abordar en los próximos capítulos es cómo se constituye el ser humano desde su biología en tanto que ser moral y qué claves ofrece ello para distinguir las competencias de la ciencia y de la ética a la hora de abordar el estudio de lo moral. Considero que el primero de los roles señalados, en función del cual la naturaleza humana compone las condiciones de posibilidad para la ética, es el que resulta de mayor interés para perseguir este objetivo. Incidiremos también en la consideración de que la naturaleza permite la fundamentación de normas y principios morales, realizando una labor crítica que será central a lo largo de toda la investigación.

La ciencia puede ofrecer una explicación y una descripción de las condiciones de posibilidad naturales que permiten la aparición de la agencia moral, que posibilita a su vez la ética. Sin embargo, a diferencia de lo que algunos neurocientíficos pretenden (Gazzaniga, 2006; Mora, 2007; Churchland, 2013; Greene), la ciencia no tiene potestad para resolver algunos de los problemas de la filosofía moral, como el de la fundamentación. La investigación naturalista de lo moral presenta límites insalvables que se harán patentes cuando el quehacer científico no pueda dar cuenta explicativa y descriptiva de aquellos aspectos que corresponde resolver a la ética, en tanto que es una disciplina normativa (Korsgaard, 2000; Cortina, 2012a; Habermas, 2006; Apel, 1985; García-Marzá, 2016; Salles, 2013, 2016, 2018).

² Cabe señalar que si bien los seres humanos pueden ser considerados dignos en virtud de la naturaleza humana, esto no excluye que solo ellos sean objetos de consideración moral. El debate sobre la extensión de la dignidad hacia otras especies animales está hoy en día en pleno auge. La cuestión incide en si lo que hace digno al ser humano es exclusivamente humano —por ejemplo la autonomía— o si bien es algo que comparte con miembros de otras especies animales —por ejemplo la sintiencia—.

Capítulo 2. Lo moral en perspectiva filogenética, ontogenética y funcional

La ética no puede reducirse al estudio naturalista de la moralidad, pero tampoco puede producirse sin las bases naturales que la posibilitan. Por ello, la labor de las ciencias empíricas en general y de las neurociencias en particular resulta de gran interés, ya que estas pueden ocuparse de investigar apropiadamente la dimensión natural de lo moral, desde el método científico. En este capítulo atenderemos al estudio de las condiciones naturales de posibilidad de la agencia moral, entendida como la facultad psicobiológica característica del agente moral, siendo esta la que le habilita como tal. Las investigaciones en este campo se producen principalmente desde tres puntos de vista, el filogenético, el ontogenético y el funcional. Su análisis revela, además, que estos tres puntos de vista están estrechamente relacionados, siendo necesaria una comprensión adecuada de cada uno de ellos para entender bien el fenómeno de la moralidad en el ser humano. El análisis crítico de este tipo de estudios permitirá comprender y aclarar algunas cuestiones metodológicas acerca de cómo la ciencia puede y debe abordar el estudio de la moralidad desde su método, que es naturalista. Todo ello sentará las bases para seguir avanzando en nuestra investigación en los siguientes capítulos.

2.1. Las condiciones naturales de posibilidad del agente moral: origen y desarrollo de la moralidad desde una perspectiva filogenética, ontogenética y funcional

La moralidad como objeto de estudio científico.

Las condiciones naturales de posibilidad para que la moralidad aparezca se componen de un conjunto de elementos cuyo estudio puede ser abordado desde las ciencias empíricas. Estos elementos son las predisposiciones psicobiológicas de prosocialidad y las facultades sociales, como la empatía o la teoría de la mente, sumadas a una serie de facultades psicológicas, habiendo alcanzado algunas de estas determinado grado de complejidad, entre ellas cabe destacar la memoria, la imaginación, el pensamiento abstracto, la razón y el lenguaje, asociadas al ejercicio de la metacognición. La concurrencia de todos esos elementos permite el ejercicio de la agencia moral, entendida como facultad característica del agente moral. Abordaremos la cuestión desde una triple perspectiva: perspectiva filogenética, sosteniendo que la moralidad surge como un subproducto evolutivo; perspectiva ontogenética, sosteniendo que la adquisición de

determinadas facultades en determinado grado de complejidad determina el paso de las conductas meramente sociales a conductas sociales que podemos calificar como morales; y perspectiva funcional, exponiendo cómo operan los elementos que componen las condiciones naturales de posibilidad de la agencia moral.

Investigar la moralidad desde el paradigma naturalista de las ciencias empíricas es posible, pero este tipo de estudios han de limitarse a investigar los elementos naturales que posibilitan la conducta moral, concretamente el sustrato neurobiológico que habilita al ser humano como agente moral. Ese sustrato biológico posibilita y condiciona lo moral en tanto que constituye su dimensión estructural, dejando abierta la posibilidad de diferentes contenidos morales, que serán contingentes, históricos y culturales (Aranguren, 1994; Cortina, 2012a; Hauser, 2008; Evers, 2010; García-Marzá, 2016). Ahora bien, no todos los aspectos relativos a la moralidad son reductibles a este paradigma. Las cuestiones relativas a la fundamentación y justificación ética de normas y valores son competencia de la filosofía moral, su estudio queda, por tanto, fuera del alcance de las ciencias empíricas.

Existen una serie de propuestas que defienden que ese sustrato biológico sirve de fundamentación para principios, valores y normas morales (Spencer, 1964, 1883; Richards, 1986; Arnhart, 1998; Singer, 1981; Flanagan; Cartwright, 2000: 328; McShea y McShea 1990, 1999; Curry, 2006: 235, 237-238; Casebeer 2003; Churchland 2013; Mora, 2007; Gazzaniga, 2006; Greene, 2013), estas no pueden resistir la crítica y tendrán que ser desechadas (Cortina, 2007, 2012a; García-Marzá, 2016; Buchanan, 2009; Apel, 1985; Habermas, 2006; Aranguren, 1994; Gould, 1999; Hume 1988, 1953; Moore, 1983; Salles, 2013, 2016, 2018; Jácomo, 2011, 2013a, 2013b, 2014; Evers, 2010; Cartwright, 2010; Teehan y diCarlo, 2004; Siurana, 2015)³. Tales propuestas presentan una pretensión de naturalismo reduccionista en relación al estudio científico de la moralidad. Este reduccionismo científicista es resultado de reconocer a la ciencia unas competencias que no le corresponden en lo que refiere a cuestiones de ética. El agente moral, en cuanto tal, es capaz de reflexionar sobre cuestiones y conceptos que no son necesariamente naturalizables. El deber moral, la justicia, el bien, la obligación moral o la justificación de normas, por mencionar algunos ejemplos, no pueden ser reducidos al ámbito de la

³ Atenderemos esta cuestión desde una perspectiva crítica a lo largo de la investigación, incidiendo especialmente en el Capítulo 6.

mera naturaleza, a pesar de que no existirían sin ciertas estructuras biológicas que son su base natural. Teniendo en cuenta esa distinción crucial atenderemos desde una perspectiva naturalista a las condiciones naturales de posibilidad de la moralidad, en tanto que base neurobiológica de esta (Ayala, 2006; Pallarés-Domínguez, 2013; Paris Albert y Comins Mingol, 2013; Salles, 2013, 2016, 2018; Cortina, 2012a; García-Marzá; Siurana, 2015).

Las conductas morales son un tipo de conducta social, esto significa que la aparición de estas en una especie precisa la previa existencia de conductas sociales en la misma. La complejización de esas conductas sociales y de las facultades que van adquiriendo y desarrollando las especies a lo largo de la evolución, así como los individuos que las forman, puede llegar a desembocar, como en el caso del ser humano, en la aparición de un tipo de conducta social que podemos calificar como moral. Trataremos de ver cómo pudo producirse en el ser humano esa transición de conductas meramente sociales a conductas sociales de tipo moral. Abordaremos la cuestión de la moralidad en los seres humanos desde una perspectiva filogenética, ontogenética y funcional comparada. Un análisis de este tipo permite comprender mejor cómo los seres humanos devienen en agentes morales en virtud de la adquisición de determinadas predisposiciones biológicas y facultades psicológicas y su ejercicio en contextos sociales.

La facultad que permite el ejercicio de la moralidad será referida en como *agencia moral*, en tanto que facultad característica del agente moral. Nuestra tesis es que la aparición de la agencia moral se ve propiciada por la concurrencia ciertas de facultades y ciertas predisposiciones biológicas, habiendo alcanzado estas cierto grado de complejidad. A las predisposiciones hacia la prosocialidad y a las facultades sociales que regulan la relación con los otros (teoría de la mente, empatía, etc.) se sumará, en el caso del ser humano, la aparición de facultades como la razón, el pensamiento abstracto, el lenguaje discursivo y la metacognición, habiendo alcanzado estas un determinado nivel de complejidad. La concurrencia de todos estos elementos permitirá esa transición de conductas sociales a conductas sociales de tipo moral.

Nuestro análisis comenzará atendiendo a las perspectivas filogenéticas de Darwin y Ayala, que respaldan esta tesis. Tras ello analizaremos cuatro modelos de desarrollo filogenético y ontogenético, en todos ellos se respalda la tesis de que los mecanismos de regulación de la conducta prosocial no bastan, se precisan facultades que permitan un

nivel metacognitivo para que el individuo se convierta en un genuino agente moral. Estos modelos son los de los primatólogos Frans de Waal y Michael Tomasello, y los de los psicólogos Piaget y Kohlberg, atenderemos también, en relación a este último, a la propuesta del desarrollo moral de Gilligan. Todo ello permitirá una comprensión funcional de las condiciones naturales de posibilidad de lo moral.

Orígenes evolutivos de la moralidad: de Darwin a Ayala.

Las conductas morales no son algo que aparezca en la evolución de forma aislada y repentina, al contrario, son el resultado de un largo proceso evolutivo, en el cual la selección natural va configurando complejas constelaciones de atributos en los miembros de las diferentes especies, en adecuación a las exigencias de supervivencia y reproducción que el medio impone. En ese proceso puede darse el caso de que una de esas constelaciones de atributos constituya las condiciones de posibilidad para que la dimensión moral aparezca.

Uno de los textos paradigmáticos y seminales en lo que concierne a la dimensión moral de la naturaleza humana es el *El Origen del Hombre*, de Darwin. En su estudio sobre los orígenes de las especies, y en concreto sobre el origen de nuestra especie, este emblemático autor advierte que la moralidad forma parte indispensable de la naturaleza humana. Las reflexiones que nos brinda supusieron un punto de inflexión en la cuestión, y a día de hoy tienen aún plena vigencia si queremos llevar a cabo una investigación como la que nos proponemos. Atender a sus escritos será nuestro primer cometido.

En el Capítulo III de *El Origen del Hombre* Darwin señala algunas de las facultades que el ser humano comparte con muchos otros animales, aun si algunas de ellas difieren en grado de complejidad en estos, entre estas facultades encontramos la memoria, la imaginación, la razón, el lenguaje, entre otras. En el Capítulo IV Darwin atiende a una facultad que –si bien participa de algunas de las facultades antes mencionadas– cobra especial importancia, ya que se da en el ser humano presentando discontinuidades muy marcadas respecto a otras especies animales, hablamos del *sentido moral*, al que acompaña la *conciencia*⁴ (Cortina, 2016, 2017). El capítulo se abre con el siguiente fragmento:

⁴ Si bien en algunos puntos Darwin identifica el sentido moral con la conciencia, dado que se dan de forma conjunta y casi indistinguible, en otros especifica que la conciencia se diferencia del sentido moral,

Acepto por completo la opinión de los escritores que sostienen que entre todas las diferencias existentes entre el hombre y los animales inferiores el sentido moral o conciencia es la más importante. El sentido moral, como dice Mackintosh, “tiene verdadera supremacía sobre todo otro principio de las acciones humanas”, y se resume en la breve pero imperiosa palabra *deber*, cuyo sentido es tan elevado. Es el más noble atributo del hombre; el que le impulsa, sin vacilaciones de ningún género, a poner en riesgo su vida por la de sus semejantes, o le mueve, tras madura deliberación, a sacrificarla en aras de una gran causa, guiado por la sola impulsión del sentimiento profundo del derecho o del deber. (Darwin, 1979: 94)

En estas palabras pueden detectarse ya algunos elementos que serán clave para descifrar en qué consiste ese sentido moral del que habla Darwin, que no es otra cosa que la facultad de la agencia moral. El papel de la razón y el sentimiento de deber son señalados por el autor como elementos centrales. Según lo entiende Darwin, este sentido moral es una de las fuerzas primordiales que guían la acción humana, y lo hace orbitando en torno a la idea de *deber*. Además, hemos de llamar la atención sobre el hecho de que Darwin dice que esto se produce “tras madura deliberación”, es decir, tras el correcto ejercicio de la razón, que ofrecería como resultado una serie de conclusiones y pautas de conducta relativas al *deber*. Al avanzar en su texto descubrimos cómo, desde una perspectiva naturalista, surge ese sentido moral, según Darwin. Este considera que la facultad del sentido moral es resultado de la complejidad y la suma de determinadas facultades intelectuales, siendo el caso, además, que todas ellas se dan en el ser humano.

[...] todo animal, cualquiera que sea su naturaleza, si está dotado de instintos sociales bien definidos, incluyendo entre éstos las afecciones paternas y filiales, inevitablemente llegaría a la adquisición del sentido moral o de la conciencia cuando sus facultades intelectuales llegasen o se aproximasen al desarrollo a que aquellas han llegado en el hombre. (Darwin, 1979: 95)

Varios elementos nuevos entran aquí en juego. En primer lugar hay que señalar el importante papel que Darwin atribuye a las emociones, que serían un elemento esencial. La mera adquisición de las facultades que permiten el sentido moral no desembocaría en este si no se diese en un animal que, como el ser humano, sea prosocial. Las emociones y facultades emocionales que predisponen al ser humano para una vida grupal compartida, entre las cuales está la simpatía por nuestros semejantes, la empatía, la vergüenza y algunas otras (Cortina, 2009, 2017; Eisenberg, 2000; Tagney, 2007), son necesarias para que el ejercicio de facultades como la razón o el lenguaje discursivo desemboque en un

cumpliendo una función en el ejercicio de este, tal función consiste en inclinar la acción en contra de la violación de aquellas pautas de conducta que el sentido moral dicta como adecuadas.

sentimiento de deber dirigido hacia los otros. Hoy en día la ciencia confirma que Darwin estaba en lo cierto, diversos estudios científicos y neurocientíficos han puesto de relieve el carácter intrínsecamente emocional de las conductas morales (Evers, 2010; Salles, 2015; Colmenares, 2005; Eisenberg, 2000; Tagny, 2007; Waal, 1997, 2007; Brosnan y Waal, 2003; Tomasello y Vaish, 2012; Churchland, 2013; Haidt, 2001; 2013; Greene; Moll, 2002; Rizzolatti, 2006; Morgado, 2010; Damasio, 2005, 2008; Cortina, 2012a, 2017; Siurana, 2011b, 2015). También la filosofía moral se hace eco de este hecho, la propuesta de una ética de la razón cordial (Cortina, 2009) y la de una neuroética de la razón cordial, que confeccionaremos a lo largo de nuestra investigación, reconocen como uno de sus pilares fundamentales la relevancia moral de las emociones y las facultades sociales.

Una segunda apreciación merece ser hecha en relación al fragmento anterior. Darwin no afirma que el sentido moral haga su aparición en el ser humano por su valor *adaptativo* (*aptativo**)⁵ y por el éxito de supervivencia que facilite en la selección natural. El desarrollo de las predisposiciones biológicas prosociales sí es *aptativo*, incluso la paradoja del altruismo biológico ha encontrado posteriormente explicaciones que logran solventar las dificultades que se planteaban inicialmente (Hamilton, 1963; Trivers, 1971; Wilson y Sober, 1994; Ayala, 2001; Kurzban, 2015). También las diversas facultades psicológicas, como la memoria, la imaginación —que implica poder hacer predicciones sobre el futuro—, la razón y el lenguaje surgen y se desarrollan en contextos de selección natural donde son *aptativas*. Sin embargo, el sentido moral no es una de aquellas predisposiciones biológicas ni una de estas facultades. Y no necesariamente surge, a diferencia de ellas, con una función *aptativa*, aunque eso no excluye la posibilidad de que una vez haya aparecido pueda desarrollarse y evolucionar *adaptativamente*. Retomaremos esta cuestión más adelante. Diversos autores, como tendremos ocasión de ver, suscriben esta tesis.

Darwin no explica cuál es el origen concreto de ese sentido moral del que habla, posiblemente porque nunca tuvo una respuesta clara al respecto. Sin embargo, atendiendo

⁵ En el próximo apartado llevaré a cabo una revisión terminológica en función de la cual el término *adaptación* será sustituido por el de *aptación*, en virtud de ello haremos uso del término *aptación* a partir de ahora como un término más riguroso que el de *adaptación* en su uso tradicional, en los casos en que usemos el término *adaptación* será en un sentido más restrictivo que el habitual, tal como establecerá la revisión terminológica indicada. *Aptación* referirá a aquello que aumenta la aptitud, sin que necesariamente se haya desarrollado para cumplir esa función, mientras que *adaptación* referirá a aquello que aumenta la aptitud pero además ha sido desarrollado específicamente para ello. Todas las adaptaciones son *aptaciones*, pero no al revés.

a su texto —como puede apreciarse en los fragmentos anteriores— podemos apreciar que sí bosqueja cuáles son las condiciones de posibilidad naturales para que ese sentido moral se produzca, a saber: el sentido moral hará su aparición en un ser cuando concurren en él una serie de predisposiciones biológicas prosociales y una serie de facultades mentales con determinado grado de desarrollo y complejidad. Si bien Darwin no logra zanjar la cuestión, algunos autores contemporáneos han tomado el relevo y se han propuesto dar una respuesta más clara al problema. Ayala, junto a Cela-Conde, toma el propio texto de Darwin para iniciar la reflexión en uno de los capítulos de *Senderos de la evolución humana*, cuyo título es “La filogénesis de la moral”. El texto pretende aclarar algunos de los elementos que tanto en *El origen del hombre* como en los textos clásicos de la sociobiología parecen quedar más difusos.

Ayala y Cela-Conde coinciden con Darwin en que una de las condiciones de posibilidad de la moral es la prosocialidad. Dado que las conductas morales son un tipo de conductas prosociales la aparición de estas ha de darse previamente o simultáneamente por necesidad. El capítulo lleva a cabo un análisis del origen evolutivo de algunas de las conductas prosociales del ser humano y cómo devienen en conductas morales, diferenciando, por ejemplo, el *altruismo biológico* del *vernáculo o moral*. El altruismo biológico se ha de entender en términos de una disposición biológica a invertir recursos propios en beneficio de otro, y no precisa de mente o de intencionalidad. Un individuo es considerado biológicamente altruista, por consiguiente, cuando invierte recursos propios en beneficiar a otros, habiendo casos en que el coste propio es elevado. A diferencia del altruismo biológico podemos encontrar un *altruismo vernáculo o moral*, que correspondería a conductas genuinamente morales. Este tipo de altruismo concierne a los motivos psicológicos para la acción, y harán del actor un ser altruista al margen de los resultados obtenidos de la acción. La presencia de mente e intencionalidad se hace precisa en este caso. Así pues, altruismo biológico y altruismo moral parecen ser dos modos bien diferenciados en los que un organismo puede invertir recursos propios en otros individuos aun si eso supone un perjuicio para sí mismo (Ayala y Cela Conde, 2001). Los autores trazarán un análisis del desarrollo evolutivo social y cognitivo del ser humano en función del cual propondrán cómo nuestros ancestros del género homo devinieron de seres meramente sociales a seres que además eran morales. Proponen un escenario en el que nuestros antepasados sufrieron la selección natural de modo tal que se potenció la adquisición y desarrollo de determinadas facultades cognitivas, como la razón y el

lenguaje. Lo valioso de la propuesta para nosotros no es hasta qué punto el escenario que proponen corresponde al que realmente se dio. Lo que nos interesa es cómo la adquisición de ciertas facultades mentales en determinado grado de complejidad marcó el paso de conductas meramente sociales a conductas sociales de tipo moral. (Ayala y Cela Conde, 2001).

Condiciones de posibilidad de la agencia moral.

Para comprender más a fondo esta cuestión hemos de remitirnos a otro texto de Ayala, en el cual este expone claramente a qué se refiere cuando habla de que el ser humano es un ser moral o ético. Este texto es un capítulo de su libro *La evolución de un evolucionista*, lleva por título “Las raíces biológicas de la moral”. En este texto Ayala trata de especificar cuáles son las condiciones necesarias y suficientes que han de darse para que un ser pueda tener conductas que calificamos de morales o éticas. Es decir, Ayala trata de establecer cuáles son las condiciones de posibilidad que permiten que un ser pueda constituirse como agente moral. Si bien Ayala expone estas condiciones de posibilidad atendiendo a tres capacidades que el agente moral precisa tener, tales capacidades implican cierto equipamiento psicobiológico que las hace posibles. Por ello, atendiendo a su propuesta podremos extraer conclusiones acerca de los elementos emocionales y cognitivos que habilitan al ser humano como agente moral. Tales elementos han de considerarse también como condiciones de posibilidad de la agencia moral. En realidad ambos enfoques atienden a las mismas condiciones, desde diferente perspectiva. Ayala atiende a las capacidades que precisa tener el agente moral, mientras que nosotros atenderemos, además, al sustrato psicobiológico que posibilita estas capacidades.

Las tres condiciones que Ayala señala son las siguientes: capacidad para anticipar las consecuencias de los propios actos, capacidad de emitir juicios de valor y capacidad de elegir entre acciones alternativas. Es decir, el agente moral tiene que ser capaz de imaginar diferentes posibilidades en lo que respecta a las diferentes acciones posibles y a las diferentes consecuencias que se deriven de estas. Además, su capacidad de hacer juicios valorativos guiará su decisión a la hora de actuar de uno u otro modo. Decisión que solo será genuinamente moral si el agente es libre en algún sentido en la toma de decisión, para lo que es preciso que tenga la capacidad de elegir entre diferentes acciones

posibles. Merece la pena volver por un momento a Darwin, pues en *El origen del hombre* se describe al agente moral de modo muy similar:

[...] llamamos morales a todas las acciones de cierta clase llevadas a cabo por un ente moral, el cual no es más que un ser capaz de comparar sus acciones o motivos pasados y futuros, y aprobarlos o desaprobados. (Darwin, 1979: 107)

Como puede apreciarse, Darwin considera que el agente moral se caracteriza por ser capaz de representarse las acciones y motivos pasados y futuros, y por la capacidad de juicio moral, cuyo ejercicio permitirá que esos motivos y acciones puedan ser aprobados o desaprobados. Darwin no menciona la tercera condición, referente al libre albedrío en el agente, en relación a esta condición tan solo podemos considerar que Darwin supone su necesidad, ya que está implícito que la deliberación y la aprobación de una acción o motivo, para ser genuinamente morales, precisan la posibilidad de escoger entre varias vías de acción diferentes.

Las tres condiciones que Ayala señala, y que también Darwin apunta, precisan la existencia de una serie de facultades. La capacidad de anticipar los propios actos precisa principalmente el uso de la imaginación, la memoria y el pensamiento abstracto, que harían posible la proyección de la situación presente hacia posibles futuros hipotéticos. En cuanto la capacidad de hacer juicios morales, esta consiste en realidad en hacer juicios sobre las razones que tenemos para actuar de un modo determinado y no de otro. Por ello, la capacidad de emitir juicios de valor es resultado de la metacognición, asociada a diferentes facultades, entre las cuales están la razón, el lenguaje, del que puede servirse esta, y el pensamiento abstracto, habiendo alcanzado todas ellas ciertos umbrales de complejidad. A ello se suman otras funciones como la memoria y la imaginación. Todos estos elementos operan en el juicio moral habitualmente, ligados a los mecanismos de prosocialidad. No obstante existen excepciones, como sucede en el caso de la psicopatía, donde pueden no operar algunos elementos de cognición social como la simpatía o la empatía, motivo por cual cabría considerar al psicópata como un caso atípico, e incluso patológico, de moralidad. La biología y la psicología señalaron hace tiempo una cuestión que la neurociencia confirma con detalle hoy en día, al ejercicio de la razón en el juicio moral se suman factores emocionales y prosociales que condicionan el juicio y que lo guían hacia ciertas direcciones (Moll, 2002; Greene; Rizzolatti, 2006; Morgado, 2010; Damasio, 2005; Damasio, 2008; Haidt, 2001; Churchland, 2013; Salles, 2015). Así pues,

como Darwin ya apuntaba, la prosocialidad del ser humano es necesaria para que el ejercicio de la razón produzca un tipo de juicio que intente resolver cómo debemos comportarnos en relación a los demás. Son las emociones prosociales las que dirigen el ejercicio de la razón hacia la búsqueda de lo mejor no solo para uno mismo sino también para los demás. Cabe señalar que una ética de la razón cordial, a la que atenderemos en futuros capítulos, retoma este hecho en el plano filosófico (Cortina, 2009). Un ser racional que no fuese prosocial posiblemente no desarrollase el sentido moral del que Darwin nos habla, como sucede en el caso, ya mencionado, de la psicopatía (Darwin, 1979; Haidt, 2001: 816; Damasio; Greene; Raine, 2000; Blair), pero podría, no obstante, emitir juicios de valor de un modo no prosocial.

Aunque el hombre, como acabamos de notar, no posee instintos especiales que le enseñen cómo debe ayudar a sus prójimos, sin embargo, existe en él ese natural impulso; y con sus altas facultades intelectuales, naturalmente se deja guiar en esto por la razón y la experiencia. (Darwin, 1979: 105)

Darwin intuyó acertadamente que los impulsos prosociales del ser humano ahondan sus raíces en las emociones. Estas sirvieron durante millones de años como mediación entre la percepción del medio y la respuesta que los organismos animales ofrecían ante este (Damasio, 2009). En muchas especies, al convertirse el grupo en el factor principal del medio al que adaptarse, las emociones evolucionaron para ofrecer una respuesta adecuada a este entorno ambiental, compuesto por los semejantes (Colmenares, 2005; Morgado, 2010). La adquisición de la razón y del pensamiento abstracto, que se produce de forma más tardía, hace que estas facultades se implementen al acervo de facultades ya existentes, cumpliendo así una función complementaria a la de las emociones y facultades prosociales, pero sin sustituirlas en sus funciones (Goldberg, 2002; Forbes y Grafman, 2010).

Resulta necesario analizar si efectivamente el ser humano cumple estas tres condiciones. En primer lugar, atenderemos a la capacidad de anticipación de acciones futuras, en segundo lugar al libre albedrío, que resulta problemático a la hora de comprobar su existencia, y en último lugar a la capacidad de hacer juicios valorativos, en la cual profundizaremos. Respecto a la anticipación de acciones futuras, tal y como ya hemos indicado anteriormente, esta es una facultad que ha sido identificada tanto en seres humanos como en otras especies animales.

En cuanto al libre albedrío, desde la perspectiva naturalista no tenemos evidencias de si el ser humano está tan solo condicionado o si está en realidad determinado por su biología, la neurociencia investiga actualmente estas cuestiones (Libet 1999; Roskies, 2010; Ortiz de Landázuri, 2016, Pallarés-Domínguez, 2013, 2016). Si bien la libertad de acción es una condición necesaria para la agencia moral no podemos más que trabajar bajo el supuesto de que el ser humano es libre en este sentido. En cualquier caso, de existir el libre albedrío, se precisarían unas complejas capacidades cognitivas, mediante las cuales fuese posible abrir una brecha en la cadena causa-efecto de la naturaleza o permitir una contingencia causal, es decir, generar al menos dos posibles efectos ante una misma causa, pudiendo así optar al menos entre dos opciones en la toma de decisiones (Evers, 2010). Por ello, dado su nivel de desarrollo cognitivo, el ser humano es uno de los animales con más probabilidades de haber alcanzado ese punto de complejidad, en caso de que fuera posible el libre albedrío. Sin embargo, hasta la fecha no se ha podido demostrar que de hecho sea así. Retomaremos esta cuestión en capítulos posteriores.

En cuanto a la capacidad de hacer juicios morales, de su adquisición depende que un ser social devenga agente moral. Ciertas facultades que existen en otras especies animales y que existieron en nuestros ancestros (razón, que puede ser también de tipo espacial e instrumental, lenguaje y pensamiento abstracto), sumadas a otras funciones como la memoria y la imaginación, permiten la aparición de una cognición compleja, que en algunos casos favorece la aparición de una conciencia no autoconsciente. Cuando algunas de estas facultades, especialmente la razón, el lenguaje y el pensamiento abstracto, cruzan ciertos umbrales de complejidad, se produce la aparición de la metacognición y de la autoconciencia. Todo ello sumado a las facultades y emociones prosociales, bajo el supuesto de que existe el libre albedrío, constituye una concurrencia de elementos que permite sentar las bases para que la agencia moral aparezca. Dicho de otro modo, esta concurrencia de elementos compone las condiciones naturales de posibilidad de la agencia moral, marcando así la transición de las conductas meramente sociales a conductas sociales de tipo moral en el ser humano. (Ver Figura 1)⁶.

⁶ Esta figura muestra las principales funciones subyacentes a cada una de las facultades a las que hemos aludido, sin que ello signifique que la totalidad de las funciones cognitivas subyacentes se encuentren reflejadas en la figura. La categoría “sin especificar/otros” permite contemplar esta cuestión, sin descartar que funciones no mencionadas en la Figura 1 estén presentes en el ejercicio de las distintas funciones indicadas. La Figura 1 es una representación que tiene como objetivo ilustrar la cuestión, pero no supone un reflejo fiel de la complejidad de las funciones cerebrales, lo cual se presenta como inabarcable e ineficiente para el objetivo que aquí se persigue.

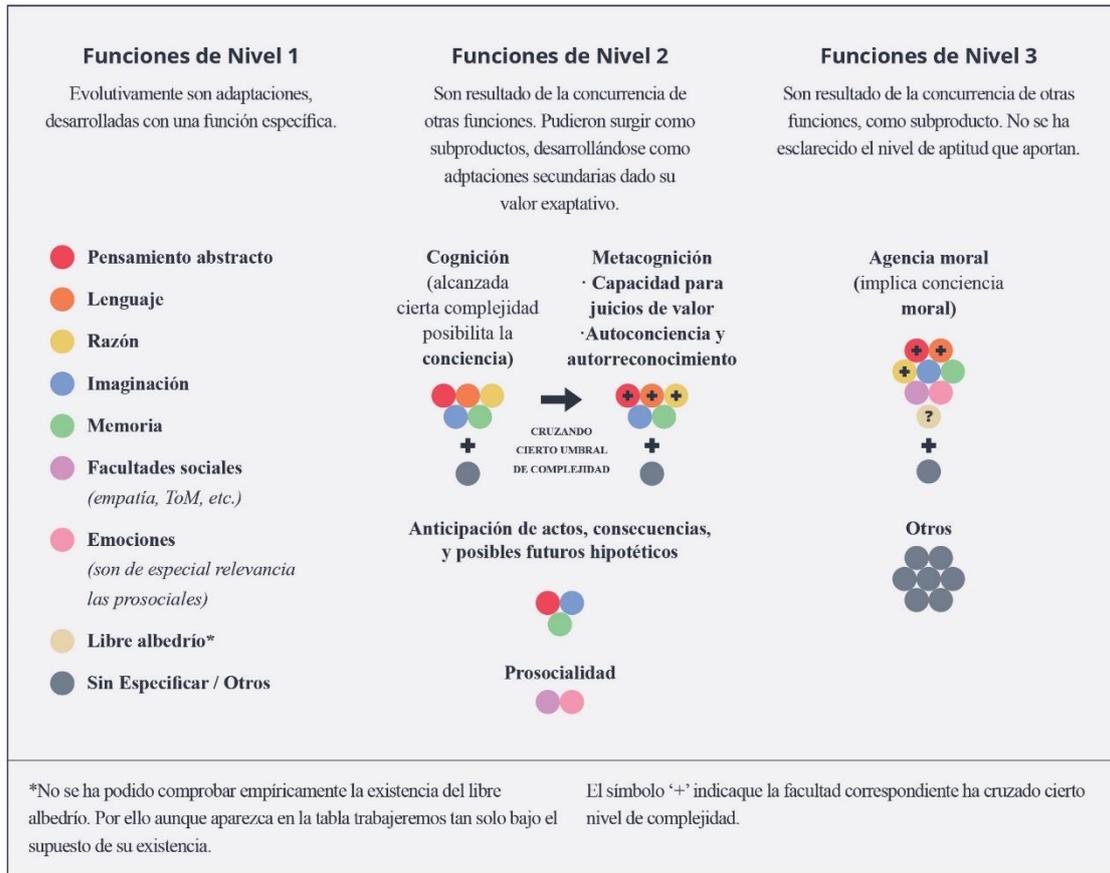


Figura 1. La agencia moral como facultad resultante de la concurrencia de otras funciones.

A las tres condiciones que Ayala expone, y que también Darwin reconoce, cabe añadir una más, la capacidad para el reconocimiento mutuo. Si bien en la propuesta de ambos el reconocimiento figura implícitamente como un elemento necesario para la moralidad, es preciso indicar explícitamente la necesidad de este. Las facultades y las emociones sociales, así como el carácter social que lo moral tiene por definición, ponen de relieve la necesidad de que el reconocimiento se produzca por parte de los agentes morales (Cortina, 2009, 2017; García-Marzá, 1992; Habermas, 2006; Apel, 1985). Un mal funcionamiento de la cognición social producirá casos atípicos de moralidad, o la ausencia de la misma, como sucede, por ejemplo, en el caso de la psicopatía.

En “Las raíces biológicas de la moral” Ayala continúa con la reflexión y ofrece algunas observaciones que son de máximo interés. Considera que la moralidad no aparece en el ser humano por ser adaptativa (aptativa*)⁷ en sí misma, algo que ya intuyó Darwin, como hemos expuesto anteriormente, y que se ve apoyado por autores como Damasio,

⁷ Ver nota 5.

Nichols y Prinz. Todos ellos sostienen que la aparición de la agencia moral no se debió a que esta favoreciese la selección natural (Damasio, 2005: 160; Prinz, 2001; Nichols, 2005). Su aparición sería una consecuencia no prevista evolutivamente, sino derivada de la adquisición y concurrencia de determinadas facultades psicológicas. Dicho de otro modo, la agencia moral surgió en un primer momento como un subproducto evolutivo (Richart, 2016).

El comportamiento ético apareció en la evolución no por ser adaptativo en sí mismo, sino como una consecuencia necesaria de las eminentes habilidades intelectuales humanas, que son un atributo promovido directamente por la evolución. (Ayala, 2006: 358)

Según plantea Ayala, la combinación de determinadas facultades cognitivas, habiendo cruzado estas un determinado umbral de complejidad, compondrían las condiciones naturales de posibilidad que convertirían a un animal social inteligente en un agente moral.

El pensamiento abstracto y la razón surgieron evolutivamente con la función de resolver problemas específicos de adaptación (aptación*)⁸ al medio. La inteligencia humana se desarrolla así de forma selectiva, configurando estas y otras facultades que van ganando poco a poco en complejidad. El ser humano se ha hecho paulatinamente portador de una rica gama de facultades cognitivas, con suficiente complejidad como para poder ser ejercitadas combinándose entre sí, lo que daría lugar a resultados y usos no previstos por la evolución. Uno de estos resultados es la agencia moral. El hecho de que no existan áreas cerebrales específicas asociadas a la moralidad, sino que esta precise del ejercicio de áreas asociadas a otras funciones (lenguaje, razón, emociones, etc.) es un indicativo de que el origen evolutivo de la moralidad radica en la concurrencia de capacidades desarrolladas primariamente para otras funciones.

Además, es probable que los sistemas que sostienen los comportamientos éticos no se hallen dedicados exclusivamente a la ética. Se dedican a la regulación biológica, la memoria, la toma de decisiones y la creatividad. Los comportamientos éticos son, por tanto, los efectos colaterales [...] de estas otras actividades. Pero no veo que haya ningún centro moral en el cerebro, y ni si quiera un sistema moral como tal. (Damasio, 2005: 160)

⁸ Ver nota 5.

La suma de tales facultades configuró al ser humano como un ser capaz de anticipar sus acciones futuras, emitir juicios morales y, además, hacerlo libremente (bajo el supuesto de que existe el libre albedrío en nuestra especie). El ser humano se convirtió así en un animal social que dotado de la suficiente complejidad cognitiva puede imaginar objetos que no están en el mundo natural, puede plantearse problemas abstractos que no son aquellos para cuya resolución emergieron las facultades con las que cuenta, puede hacerse preguntas acerca del mundo, puede imaginar otros mundos posibles, preguntarse si entre ellos hay alguno que sea moralmente mejor que otro, buscando cómo lograr su realización a través de sus propias acciones y las de los demás, lo cual precisa justificar normativamente tales acciones, así ciertas normas y códigos de conducta que velan por su cumplimiento. Así, el conjunto de facultades anteriormente descrito, habiendo alcanzado el nivel de complejidad preciso, se abre un espacio de reflexión donde el ser humano puede preguntarse cómo *debe ser* el mundo, ya no solo *qué es* o *cómo es*. Comprender la naturaleza de este salto cognitivo es de máxima importancia para dirimir qué es la ética y hasta dónde puede llegar la ciencia en el estudio de lo moral. La tesis de que la agencia moral surge como un subproducto evolutivo dada la concurrencia de otros elementos cognitivos, a su vez, podría ofrecer una explicación de otros elementos presentes en el ser humano, como el hecho religioso o del hecho artístico (ver Figura 1, objeto “otros”). Dirimir si esta explicación ofrece la mejor explicación para tales hechos corresponderá a ulteriores investigaciones.

Modelos filogenéticos y ontogenéticos del desarrollo moral: F. de Waal, M. Tomasello, J. Piaget y L. Kohlberg.

Para apoyar la tesis que hemos defendido desde una perspectiva evolutiva, atenderemos ahora a cuatro modelos del desarrollo moral, ampliando el enfoque también desde el punto de vista ontogenético. Para ello atenderemos a las propuestas de los primatólogos F. de Waal y M. Tomasello, así como a las de los psicólogos J. Piaget y L. Kohlberg. Es significativo que todos ellos coincidan con Ayala en que precisamente la adquisición de las facultades que permiten la metacognición, asociada al pensamiento abstracto, la razón y el lenguaje discursivo, marca el punto de inflexión en el que un animal social va configurándose como agente moral. A pesar de las distintas corrientes o perspectivas en las que estos autores se encuadran, creo que es significativo el modo en que todos ellos reconocen que en la configuración del agente moral existe una gradación

de sus capacidades, las cuales presentan sus máximas posibilidades al adquirir ciertas facultades y alcanzar estas ciertos niveles de complejidad. El análisis de los diferentes modelos que estos autores proponen permitirá comprender al agente moral en un proceso gradado de adquisición de capacidades, tanto filogenética como ontogenéticamente, en lugar de considerar al agente como ya realizado desde su nacimiento.

Frans de Waal ha llevado a cabo en los últimos años investigaciones comparadas que son de gran interés en torno a los orígenes de lo moral (Waal, 1997, 2007). Este diferencia tres niveles de elementos que configuran la dimensión moral en el ser humano, algunos de los cuales se dan en otras especies. Indica la existencia de un primer nivel correspondiente a los motivos emotivo-psicológicos de la acción, en el cual encontramos factores como la empatía, la reciprocidad, la atribución, la disposición a la resolución de conflictos, el sentimiento de justicia e injusticia –entendidos, respectivamente, como satisfacción e insatisfacción de expectativas en relación a recursos invertidos–, la propensión a establecer relaciones armónicas, etc. En un segundo nivel sitúa la presión social, consistente en la insistencia por parte de los miembros del grupo en que cada uno de los demás individuos se comporte de forma favorable para la vida cooperativa grupal. La recompensa, el castigo y la construcción de la reputación son clave en este nivel. La normatividad también se sitúa aquí, pero difiere mucho entre especies. Mientras las pautas de conducta aprendidas se dan de forma implícita en la vida grupal en la mayoría de animales sociales, en el ser humano puede alcanzar un carácter explícito, pudiendo, incluso, producirse la positivación de esas normas. Con ello llegamos al tercer nivel. Mientras que los dos niveles anteriores son compartidos por el ser humano, de uno u otro modo, con otras especies animales, el tercer nivel se da tan solo en nuestra especie. En él se sitúa la reflexión racional que permite hacer juicios éticos. Elementos propios de este nivel son la interiorización de las necesidades y objetivos de los demás, la autorreflexión, el pensamiento abstracto, el razonamiento lógico y lingüístico, y la metacognición. (Waal, 2007).

Michael Tomasello, por su parte, desde una perspectiva comparada, también ha ofrecido interesantes aportaciones sobre la socialidad en primates, especialmente en chimpancés, y sobre la moralidad en humanos. En su artículo “Origins of Human Cooperation and Morality” (Tomasello y Vaish, 2013) ofrece un doble análisis, filogenético y ontogenético, del surgimiento de la moralidad en los seres humanos. La

propuesta de Tomasello es que en ambos casos el desarrollo de la moral sucede en dos pasos, un primer paso en el que la conducta se dirige a individuos concretos del entorno del sujeto y un segundo paso en el cual la interiorización del deber adquiere un carácter generalizado e impersonal.

En ambos casos, el primer paso en la secuencia es la colaboración mutua y las interacciones prosocialmente motivadas con otros individuos específicos, y el segundo paso es la moralidad abstracta, de agente neutral, basada en normas de individuos que viven en mundos culturales de escala mayor llenos de convenciones, normas e instituciones impersonales y mutuamente conocidas. (Tomasello, 2012: 232)

El primer paso que Tomasello señala presenta elementos que coinciden en muchos aspectos con los dos primeros niveles de Frans de Waal. Ambos sitúan en estos niveles las emociones y facultades sociales, las predisposiciones biológicas favorables a la vida grupal y la influencia del propio grupo. El segundo paso del que habla el fragmento anterior tiene mucho que ver con las facultades que Frans de Waal sitúa en el tercer nivel. La moralidad abstracta de carácter neutral e impersonal basada en normas, de la que habla Tomasello, es solo posible en un genuino agente moral que sea titular de facultades como la razón, el pensamiento abstracto y el lenguaje discursivo, alcanzando estas cierto grado de complejidad, como también indica Frans de Waal.

Un ejemplo de la coincidencia entre ambos autores lo ilustra la cuestión de la justicia. Según Frans de Waal el sentimiento de justicia, asociado a la inversión de recursos y la satisfacción de expectativas, se produce primariamente de forma egocéntrica y endogámica, según muestran los estudios con primates (Brosnan y Waal, 2003), tan solo con el ejercicio de las facultades cognitivas del tercer nivel pueden generalizarse o universalizarse de forma impersonal y desinteresada las expectativas de justicia, solo en este nivel podríamos hablar genuinamente de justicia en sentido moral. Este punto de vista es compartido por Tomasello en la medida en que considera que las pautas que rigen la colaboración entre individuos solo pueden adquirir un carácter impersonal, neutral y normativamente sistematizado con la adquisición de tales facultades cognitivas. Tomasello señala, además, que existe una diferencia clave entre las expectativas conductuales que configuran las sociedades de grandes simios y las humanas. Mientras que aquellas están gobernadas por la expectativa de lo que otros harán, estas últimas sin embargo lo están por la expectativa de lo que los otros deben hacer.

De muchos modos, la característica más distintiva de la organización social humana es su estructura normativa. Los seres humanos no solo tienen expectativas estadísticas sobre lo que otros *harán* —que todos los simios tienen— tienen además expectativas normativas sobre lo que otros *deben* hacer. (Tomasello, 2012: 238)

¿Qué es lo que permite esa transición de las conductas meramente sociales a conductas sociales de tipo moral? ¿Qué convierte las expectativas sobre lo que otro hará en expectativas sobre lo que debe hacer? Tomasello coincide con Frans de Waal y con Ayala en este punto. El desarrollo de determinadas facultades asociadas a la metacognición, como el pensamiento abstracto, el lenguaje discursivo y la razón, posibilitan la transición de ese primer paso dirigido a individuos específicos a ese segundo paso marcado por un sistema de normas impersonal, abstracto y neutral.

Una cuestión que surge es qué contribuye al cambio ontogenético de una moralidad de segunda persona a una basada en la norma. Una parte de la respuesta ciertamente radica en el desarrollo social-cognitivo. Para involucrarse en una moralidad basada en la norma, los niños deben pasar de ver a los individuos y las interacciones sociales puramente en términos interpersonales a adicionalmente ver todos los individuos desde una perspectiva de agente neutral o de ojo de pájaro (la “vista desde ninguna parte”; Nagel 1986). También deben desarrollar la capacidad de verse a sí mismos como individuos al igual que todos los otros individuos, para evaluar su propio comportamiento, y para entender que otros los evalúen en el mismo modo que ellos evalúan a otros. (Tomasello, 2012: 250)

Si bien este fragmento refiere al desarrollo ontogenético, el paralelismo que Tomasello establece con el desarrollo filogenético de lo moral en humanos nos permite afirmar que el surgimiento de las facultades anteriormente indicadas —tanto en la evolución como en la ontogénesis— es lo que permite la transición de la mera socialidad a la moralidad, propiciada por la emergencia de la agencia moral.

Para profundizar más en el enfoque ontogenético merece la pena atender a dos autores clásicos en la cuestión, Piaget y Kohlberg, quienes ofrecen un modelo del desarrollo moral basado en diferentes niveles. Ambos modelos contemplan la existencia de un nivel de desarrollo que sería exclusivo del agente moral, cuyo funcionamiento requiere la participación de las facultades antes mencionadas. Estos dos modelos no solo son compatibles con el modelo filogenético y ontogenético de Tomasello y con el filogenético de Frans de Waal, sino que permiten respaldarlos.

El modelo ontogenético de Piaget (1987) establece tres niveles en el desarrollo moral. El primero de los niveles se caracteriza por una etapa premoral, en la que no existe una concepción definida de la moralidad (0-5 años). El segundo de los niveles se caracteriza por una etapa de moralidad heterónoma, donde el establecimiento de las normas no depende del individuo. Estas se asumen de forma acrítica, las consecuencias de la violación de la norma regulan su adquisición, sin importar la intencionalidad de esta o su justificación (5-10). Pueden encontrarse coincidencias de esta etapa con los niveles 1 y 2 de Frans de Waal y con el primer paso de Tomasello. Por último, existe un nivel en el que encontramos una etapa de moralidad autónoma, basada en la equidad, que precisa el ejercicio de las facultades mencionadas anteriormente, a saber, razón, pensamiento abstracto, lenguaje discursivo y metacognición.

El modelo de Kohlberg es cercano al de Piaget, pero presenta algunas diferencias sustanciales. De hecho Kohlberg toma algunos modelos como referencia, entre ellos el de Piaget, demarcándose respecto a ellos, y ejerciendo una crítica en los aspectos que considera pertinente (Kohlberg, 1992: 72-95). Su modelo caracteriza el desarrollo moral en tres niveles, dividiendo cada uno de ellos en dos estadios (Kohlberg, 1992: 72-95). El primero de ellos es el nivel pre-convencional (4-10 años), en él encontramos los Estadios 1 y 2. El Estadio 1 está basado en la obediencia y el castigo. En él el individuo se comporta de forma egocéntrica, el respeto a las reglas se ve sujeto a la evitación del castigo y a la obtención de recompensas. El Estadio 2 está basado en el individualismo, la finalidad instrumental y el intercambio. El respeto de las reglas se ve sujeto a los intereses individuales, reconociendo la existencia de los intereses de otros, lo cual hace que las relaciones con estos tengan un valor instrumental, mediando el intercambio y el acuerdo si es necesario.

El segundo de los niveles es el convencional (10 años-adolescencia), compuesto por los Estadios 3 y 4. El Estadio 3 está basado en la reciprocidad y la conformación interpersonal. En él entra en juego la reputación, el respeto de las reglas viene sujeto a la necesidad de reconocimiento social en contextos de reciprocidad. El Estadio 4 está basado en el mantenimiento del orden social. El seguimiento de las reglas se halla sujeto al valor que se le otorga a ese orden como garante del buen funcionamiento social.

El tercer nivel es el pos-convencional, en el que encontramos criterios globales de justicia, estando dividido en los Estadios 5 y 6. El Estadio 5 está basado en el contrato

social. Los sistemas de normas y valores son reconocidos como contingentes, y el mantenimiento del contrato garantiza salvaguardar los principios y normas del sistema. El Estadio 6 está basado en el reconocimiento de principios éticos universales, fundamentados racionalmente, respecto a los cuales se adquiere un compromiso. El individuo adquiere en este estadio un conjunto de valores y principios de forma autónoma.

Los niveles pre-convencional y convencional de Kohlberg presentan algunas coincidencias con los niveles 1 y 2 de Frans de Waal, con el primer paso de Tomasello y con la moralidad heterónoma de Piaget, si bien pueden identificarse algunos matices que los diferencian. En cuanto al nivel pos-convencional el papel de la razón, el pensamiento abstracto, el lenguaje discursivo y la metacognición se hacen imprescindibles, y son, como el Estadio 6 indica, el elemento que permite configurar al individuo como un genuino agente moral. Los elementos a los que se ve sujeto el seguimiento de las normas en los Estadios del 1 al 4 podrían no presentar estas facultades en su máximo nivel de complejidad, sin embargo en los Estadios 5 y 6 tal cosa no es posible. La caracterización que el propio Kohlberg hace del Estadio 6 pone en evidencia la necesidad de que el individuo se rija por unos criterios de racionalidad y de lógica, que solo pueden ser ejercidos metacognitivamente:

Estadio 6: Orientación de conciencia o principio. Orientación no solo hacia reglas sociales ordenadas sino a principios de elección que requieren la llamada a una consistencia y universalidad lógica. (Kohlberg, 1992: 80)

El modelo de desarrollo moral de Kohlberg, de corte racionalista, es modificado y ampliado por Carol Gilligan (2003), en cuyos estudios se introducen elementos emocionales y asociados al cuidado, subrayando así el importante papel de lo emocional en el agente moral. Los estudios de Gilligan son una aportación valiosa para una perspectiva cordial de la moralidad, permitiendo reconfigurar las propuestas exclusivamente racionalistas en propuestas cordiales (Cortina, 2009), ampliando y modificando aquellas de modo que se evite excluir los componentes emocionales de la conducta moral. Estos intervienen de forma necesaria en el ejercicio de la moralidad, ya que la razón integra es cordial, produciéndose siempre de forma encarnada, y no como una razón pura.

Así pues, desde los modelos de desarrollo moral de estos autores se ve respaldada la tesis de que la moralidad precisa la presencia no solo de predisposiciones hacia la

prosocialidad y de mecanismos de regulación de la conducta en el grupo, sino también de la razón y del pensamiento abstracto, asociados al ejercicio del lenguaje discursivo, elementos que operan conjuntamente asociados a la metacognición, todo ello habiendo alcanzado determinado umbral de complejidad (ver Figura 1 y Tabla 1).

COINCIDENCIAS Y EQUIVALENCIAS	Waal	Tomasello	Piaget	Kohlberg
	Modelos filogenéticos		Modelos ontogenéticos	
Pueden identificarse coincidencias entre los modelos, aunque no hay una equivalencia completa.	Nivel 1: motivos psicológicos de la acción.	Paso 1: conducta egocéntrica y endogámica, dirigida a individuos del entorno del sujeto.	Nivel pre-moral: no existe una concepción definida de la moralidad.	Nivel pre-convencional: · E1: obediencia, castigo. · E2: individualidad, intercambio.
	Nivel 2: presión social.		Moral heterónoma: normas impuestas, regulado por castigo y recompensa.	Nivel convencional: · E3: relaciones interpersonales. · E4: orden social.
Paso de conductas sociales a conductas morales. Existe coincidencia en las facultades necesarias para esa transición: razón, pensamiento abstracto, lenguaje discursivo, metacognición. (Estas se sumarían a las predisposiciones de prosocialidad previamente existentes).	Nivel 3: juicio moral.	Paso 2: interiorización del deber con carácter generalizado e impersonal.	Moral autónoma: basada en equidad.	Nivel pos-convencional: · E5: contrato social. · E6: principios éticos, universalidad, autonomía.

Tabla 1. Modelos filogenéticos y ontogenéticos del desarrollo moral, coincidencias y equivalencias.

Crítica a los modelos no metacognitivos de lo moral.

En los apartados anteriores hemos visto cómo los diferentes autores a los que hemos atendido definían al ser humano como un animal cuyas facultades y la complejidad de las mismas abrían ante sí el problema de lo moralmente normativo en contextos sociales, constituyéndose como agente moral. Según esta concepción de lo moral, el ser humano sería el único animal genuinamente moral. Sin embargo, otros autores consideran que las conductas morales no son exclusivas del ser humano, sino que se dan en cierto modo en otros animales. Si definimos lo moral como una dimensión conductual de resolución de problemas sociales, donde lo normativo se plantea como un problema explícito a la hora de regular la conducta, al que se trata de dar respuesta desde la autorreflexión y el asentimiento ante normas, entonces solo los seres humanos serán genuinamente morales. Sin embargo, si consideramos lo moral en sentido más amplio, muchos animales no humanos sí serían morales. En este sentido el problema parece terminológico.

Algunas posturas plantean que un animal puede considerarse moral en un sentido amplio del término, en la medida en que presente una dimensión de resolución de problemas sociales donde lo normativo esté presente de forma implícita en la vida grupal, guiando la acción en el entorno social, si bien no necesariamente se puede explicitar y positivar lingüísticamente (Waal, 1997, 2007; Churchland, 2013). Si de este modo llamásemos morales ciertas conductas de este tipo, no serían morales en el mismo sentido en que lo son en el ser humano. Suponiendo que hubiese conductas morales en su sentido más etimológico, el de mores, en otras especies animales, estas serían distinguibles de la moralidad humana en tanto que esta tendría su propia especificidad. Solo los seres humanos presentamos el nivel metacognitivo descrito en apartados anteriores. Y solo los humanos, en virtud de ello, son capaces de reflexionar éticamente sobre lo moral, teniendo en cuenta que la ética es la reflexión filosófica sobre las cuestiones morales. Tan solo el ser humano es moral en este sentido, o si se prefiere, tan solo el ser humano puede ser ético.

Habría que preguntarse, entonces, si tiene sentido calificar de morales cierto tipo de conductas sociales desprovistas de la dimensión ética. Churchland sostiene que sí, atendamos a su propuesta para considerar los problemas relativos a esta cuestión. La autora sostiene que en otras especies de animales existe la moralidad, presentando esta su

propia especificidad también entre esas especies. Por ejemplo, en chimpancés y delfines encontraríamos diferentes modos de darse lo moral. En este sentido sería un error considerar que solo el ser humano es moral aduciendo que tiene formas específicas de moralidad. Lo adecuado sería decir que entre las distintas formas de lo moral, que se presenta de forma diferentes en diferentes especies, la del ser humano tiene unas características que la definen en contraste con el resto, siendo resultado del papel que ciertas facultades cognitivas, con determinado nivel de complejidad, juegan en la resolución de problemas sociales. El problema parece ser más terminológico que ontológico, que los animales tengan o no conductas morales dependerá de qué entendamos por conducta moral. Lo que parece ser indiscutible, en cualquier caso, es que la moralidad humana permite que exista la dimensión ética, en tanto que esta precisa una complejidad en la metacognición, la razón y el lenguaje que no está presente en otras especies. Entendiendo así la cuestión, ser moral no sería algo exclusivamente humano, pero sí ser ético.

La tesis que Churchland defiende es que aquello que los humanos llamamos ética o moralidad puede definirse atendiendo a cuatro dimensiones que se configuran a lo largo del proceso evolutivo, que interactúan a nivel cerebral y que componen una estructura de conducta social:

La hipótesis predominante es que lo que nosotros, los humanos, llamamos ética o moralidad es una estructura de conducta social en cuatro dimensiones que viene determinada por la interrelación de distintos procesos cerebrales: (1) El cuidado o la atención a los demás (enraizado en el apego a nuestros familiares y la preocupación por su bienestar), (2) el reconocimiento de los estados psicológicos de los demás, (basado en las ventajas de predecir la conducta de terceros), (3) la resolución de problemas en un contexto social (por ejemplo como deberíamos distribuir los bienes cuando son escasos, cómo resolver disputas territoriales o cómo deberíamos castigar a los sinvergüenzas) y (4) el aprendizaje de prácticas sociales (mediante un refuerzo positivo y negativo, por imitación, por ensayo y error, por diversos condicionamientos y por analogía). (Churchland, 2013: 19)

Esta descripción de lo moral se ajusta a los niveles más básicos de desarrollo de la agencia moral, descritos en los apartados anteriores, y no toma en consideración el nivel en el que la metacognición permite la apertura del espacio de la ética, entendida como reflexión sobre lo moral. Sin embargo, Churchland refiere a ello como moral y como ética, de forma sinónímica, lo cual significa que la autora no hace la distinción que

hemos señalado anteriormente respecto a la especificidad de la moralidad humana, en caso de que consideremos que hay otras especies con algún tipo de conducta moral, en sentido amplio. Si bien Churchland puede defender este concepto amplio de moralidad, extensible a otras especies, no debería calificarlo de ético, por las razones expuestas. Asimismo, debería especificar los aspectos propios de la moralidad en el ser humano. Todo ello denota que la autora está usando un concepto de ética y de moral poco riguroso que no es aceptable en los términos en que lo presenta.

De hecho, llega incluso a negar que el papel de la razón sea relevante para describir lo moral. La autora denuncia las actitudes excesivamente racionalistas en lo que concierne a la cuestión de dirimir si la dimensión moral es exclusiva o no de nuestra especie, resulta especialmente interesante el hecho de que Churchland arremete contra la definición de moralidad propuesta por Korsgaard.

[...] haríamos bien en evitar una suposición según la cual sólo los seres humanos gozan de «verdadera» moralidad; [...]. En parte las aportaciones a esta cuestión dependen del rigor con el que utilicemos nuestras palabras. No hay ningún «zar del significado» cuya opinión rijan el uso de las palabras. Si definimos las palabras de un modo en que la verdadera moralidad requiera un lenguaje y unas normas formuladas en términos lingüísticos, entonces sí, podemos deducir que sólo los humanos gozan de verdadera moralidad. Pero ¿qué avances obtenemos con esta normativa semántica? Y, en cualquier caso, ¿por qué definir la «verdadera moralidad» como algo que requiera un lenguaje? Algunos autores, como la filósofa contemporánea Christine Korsgaard, se ciñen a un argumento muy distinto: sólo los seres humanos son genuinamente racionales, la moralidad depende de la racionalidad, y por tanto los animales no humanos no son morales. (Churchland, 2013: 37)

Churchland trata de defender su postura relativizando el problema terminológico, pero no considera la posibilidad de que, incluso aceptando un concepto amplio de moralidad, compartido con otras especies, el ser humano presente una especificidad en su modo de ser moral, puesto que sólo nuestra especie puede ser ética, entendiendo lo ético como la reflexión sobre lo moral.

Las investigaciones etológicas en otras especies de animales sociales, especialmente en primates, revelan que muchas conductas y facultades humanas vinculadas a la moralidad presentan correlatos en otras especies. Independientemente de si reconocemos como morales o no tales conductas y facultades, el hecho de que sean relevantes para comprender nuestra propia conducta moral es un indicador de que guardan

estrechos lazos con esta. Tal y como hemos expuesto en el apartado anterior, la concurrencia de una serie de facultades y de elementos cognitivos y emocionales, habiendo alcanzado en algunos casos cierto nivel de complejidad, permiten el surgimiento de la agencia moral, entendida como facultad psicológica que permite lo moral. Desde esta perspectiva, podemos reconocer en algunas especies animales algunos o todos estos elementos, en diferentes grados de complejidad, sin que eso signifique que cumplan las mismas condiciones para la moralidad que se dan en el ser humano.

Las formas de altruismo moral de nuestra especie derivan y participan de formas de altruismo biológico, presentes en nuestros ancestros y en otras especies (Ayala y Cela Conde, 2001). Cabe, pues, considerar que existe o ha existido un número considerable de especies que presentan elementos cognitivos que marcan la transición de ese altruismo meramente biológico hacia uno moral o vernáculo —en términos de Ayala—. Estas especies presentan elementos cognitivos y sociales que son la antesala de lo que configurará la conducta moral de nuestra especie. Llamar *building blocks* (Joyce), protomoral (Zubiri; Aranguren; Haidt, 2001: 18-19) o moral (Churchland; Waal) a este tipo de elementos es algo puramente convencional, no obstante es cierto que las dos primeras opciones inducen menos confusión que la última opción. Ayala y Cela Conde aluden a las similitudes entre nuestra especie y las de algunos primates que son cercanos evolutivamente, reconocen que ha de existir cierto grado de continuidad entre algunas especies de simios, y reconocen la presencia de conductas sociales tan complejas como evitar obligaciones (Ayala y Cela Conde, 2001: 526), engañar y aprovechar las circunstancias a su favor (Ayala y Cela Conde, 2001: 534). Queda implícito en tales afirmaciones el reconocimiento en otras especies de la capacidad de predecir la conducta de otros y de conocer qué conductas son aceptadas y cuáles no dentro del grupo. La presencia de estas capacidades en otras especies son un indicador de que la conducta moral tal y como se da en nuestra especie no es algo que emerja de forma repentina, sino que se da como consecuencia de elementos cognitivos surgidos evolutivamente con anterioridad. La complejidad en el desarrollo de las facultades humanas es lo que posibilita la aparición de las complejas formas de conducta moral de nuestra especie. Los autores señalan la dificultad de determinar cuándo dos conductas similares responden a motivos psicológicos y/o instintivos diferentes. Calificar de heroico un acto altruista de un babuino o un chimpancé parece ser algo más difícil de justificar que hacerlo respecto a un humano. Esto no significa que esos primates no sean capaces de realizar un acto

altruista de ese tipo, lo que significa es que encontramos serias barreras epistemológicas a la hora de conocer si tal cosa es posible. En cualquier caso, los miembros de estas especies no serían éticos, en el sentido propio del término. Deberíamos hacernos la pregunta de si pueden existir conductas morales en seres que no tienen la capacidad de ser éticos.

Las especies que presenten algunas condiciones necesarias para la agencia moral, pero no suficientes, no pueden ser consideradas morales en el sentido en que lo es el ser humano. Algunos autores han hablado de ellas como poseedoras de *elementos de construcción de lo moral (building blocks)* (Joyce), otros han referido a esta antesala de lo moral como *protomoral* (Zubiri, Aranguren). Si convenimos en llamarlo *moral*, definiendo este tipo de moral de forma más amplia y excluyendo los elementos que son propios de la especificidad moral humana, como Churchland sugiere, entonces sí, podremos calificar de morales a otras especies. Churchland tiene razón en que el problema es terminológico, podemos calificar de morales ciertas conductas de otras especies si así lo deseamos, pero deberemos de mantener un término para referir a lo específicamente humano en la moralidad, ya que el modo en que nuestra especie es moral es cualitativamente diferente al modo en que otras especies podrían serlo en un sentido amplio o débil del término. Tan solo el ser humano es ético, como ya hemos afirmado, aunque convengamos en llamar morales cierto tipo de conductas en otras especies. En cualquier caso, para evitar una ambigüedad y una vaguedad terminológica, considero más adecuado hablar de *protomoral* y de *building blocks* en relación a otras especies de animales sociales, conservando los términos *moral* y *ética* para referir a la conducta humana.

2.2. Profundizando en una comprensión de los orígenes de la moralidad

La agencia moral, entendida como la facultad que permite la moralidad, aparece gracias a las estructuras psicobiológicas que componen sus condiciones de posibilidad. Trataremos de comprender su origen evolutivo atendiendo a la posibilidad de que la agencia moral surgiese como un subproducto de la evolución, que no aparece con una función concreta, y que sería resultado de la concurrencia de otros elementos que surgieron y se desarrollaron adaptativamente, tal y como hemos expuesto en el apartado anterior. Tras ello haremos algunas consideraciones respecto a las competencias de las ciencias empíricas y de la ética en relación a lo moral.

Revisión terminológica sobre factores evolutivos relativos a la aptitud.

La hegemonía del programa adaptacionista en la biología evolutiva ha dejado ciertos factores evolutivos importantes en la sombra. Una aclaración terminológica relativa a estos factores permitirá analizar con rigor la naturaleza evolutiva de la agencia moral, así como las diferentes hipótesis sobre su origen. Entre ellas la tesis que defenderemos es que la agencia moral surge como subproducto evolutivo, no siendo necesariamente aptativa.

A finales de los 70 (Gould y Lewontin, 1979) y a principios de los 80 (Gould y Vrba, 1982) vieron la luz una serie de artículos que cuestionaban la suficiencia del concepto de *adaptación* (*adaptation*) en relación a la evolución de las especies⁹, introduciendo conceptos nuevos como *exaptación* (*exaptation*), *aptación* (*aptation*), *enjuta* o *pechina* —según la traducción— (*spandrel*), reconociendo además la relevancia de los *subproductos* evolutivos (*by-products*).

En primer lugar es conveniente definir qué entendemos por *aptitud* en biología evolutiva: la *aptitud* (*fitness*) refiere al grado de adecuación de una especie a su medio, cuanto mayor sea la aptitud se obtiene un mayor éxito evolutivo en la selección natural, sujeto a dos factores principalmente, la supervivencia y la reproducción.

Etimológicamente *adaptación* significa “apto” (*aptus*) “para” (*ad*). Han predominado dos usos del término, uno de ellos al no ser riguroso ha generado una serie de errores y confusiones. Los dos usos son los siguientes:

- *Uso riguroso*: una característica es considerada una adaptación únicamente si mejora la aptitud y ha sido desarrollada por la selección natural para (*ad*) desempeñar la función que actualmente desempeña (Williams, 1966; Vrba y Gould, 1979).
- *Uso laxo*: una característica es considerada una adaptación si supone una mejora de la aptitud (*fitness*), independientemente de su origen filogenético. Yerra al atender solo a la utilidad actual de una característica, desatendiendo su origen evolutivo (Bock, 1980).

⁹ Para una crítica detallada al programa adaptacionista ver: Gould y Lewontin, 1979.

Tomaremos el uso riguroso del término. Hay casos en los que una adaptación o un subproducto de esta puede tener efectos colaterales que mejoren la aptitud. En este caso no se produce una adaptación en rigor. Para estos casos, dicen Gould y Vrba, es preciso acuñar un término del que carece hasta entonces la biología evolutiva. Consideran que es apropiado el término de *exaptación* (*ex aptus*, apto desde).

Sugerimos que estas características, evolucionadas para otros usos (o para ninguna función en absoluto), y más tarde “cooptadas” para su papel actual, sean llamadas *ex-aptaciones*. [...] Son aptas para su papel actual, por lo tanto *aptus*, pero no fueron diseñadas para ello, y no son por consiguiente *ad aptus*, o empujadas hacia la aptitud. Deben su aptitud a características presentes por otras razones, y son por consiguiente *aptas (aptus) por razón de (ex) su forma, o ex aptus*. (Gould y Vrba, 1982: 6)

Una definición sumaria del término sería la siguiente:

- *Exaptación*: Una característica es considerada una exaptación cuando aumenta la aptitud del individuo, sin que esa característica haya surgido en la selección natural para el fin que ahora desempeña.

La exaptación puede darse por diferentes circunstancias:

- a) Una característica cumple una función por adaptación, y además tiene un efecto que mejora la aptitud por exaptación.

Ej. Las plumas en las extremidades de algunos terópodos (dinosaurios ancestros de las actuales aves) por adaptación permitían una ventaja en la caza de insectos y en la termorregulación, por exaptación resultaron útiles para vuelos cortos, siendo posteriormente desarrolladas como adaptación para el vuelo. Las adaptaciones con origen exaptativo se llaman *adaptaciones secundarias*. Muchas de las actuales adaptaciones tuvieron un origen exaptativo.

- b) Como subproducto evolutivo, que tiene el efecto casual de mejorar la aptitud.

Ej. La dominancia de las hembras de hiena moteada supuso un aumento de la testosterona, este produjo como subproducto una morfología masculina en los genitales femeninos, lo cual tuvo el efecto de ser útil para ciertos rituales

afiliativos de la especie, en los que dos individuos exponen mutuamente sus genitales.

En ambos casos, el de exaptación y el de adaptación, se produce una mejora de la aptitud del individuo. Dado que en ambos casos la aptitud aumenta, Gould y Vrba proponen un concepto que engloba a ambos, el de *aptación*. Este adopta la definición del uso laxo que tradicionalmente se había hecho del término adaptación (defendido por Bock y otros):

- *Aptación*: Una característica es considerada una aptación si supone una mejora de la aptitud, independientemente de su origen filogenético. La adaptación y la exaptación son sus modos.

Gould y Vrba sugieren explícitamente que este término ha de sustituir al de adaptación en su uso laxo (Bock), por no ser este riguroso.

Algunos colegas han dicho que prefieren la definición más general de Bock porque es más fácilmente operacional. [...]

A esto respondemos que no estamos tratando de dismantelar el concepto de Bock. Nos limitamos a argumentar que este debería ser llamado aptación (con la adaptación y la exaptación como sus modos). (Gould y Vrba, 1982: 7)

Hemos manejado el término de *subproducto* (*by-product*), pasemos a definirlo. Un subproducto evolutivo es una característica o estructura que surge como resultado colateral en el desarrollo de adaptaciones. Los subproductos que resulten aptativos pueden ser cooptados como una exaptación. La cooptación es el proceso mediante el cual una característica que resulta aptativa de forma colateral es seleccionada para un uso distinto al que desempeña por adaptación. Sirva el ejemplo de la hiena moteada. Una traducción más rigurosa del término *by-product* es la de *coproducto*, ya que refleja mejor el hecho de la aparición conjunta de dos o más características, así como la aparición de dos o más usos para una misma característica, ajustándose así de forma más rigurosa al significado del término *by*. Podemos manejar indistintamente ambos términos, sin embargo a lo largo de la investigación optaremos por el de *subproducto*, ya que es el más utilizado en biología evolutiva.

Un tipo de subproducto es el llamado *enjuta* o *pechina* —según traducción— (*spandrel*). El concepto es acuñado por Gould y Lewontin como analogía con la estructura arquitectónica del mismo nombre. En arquitectura la enjuta ocupa el espacio que queda como resultado de superponer dos estructuras que tienen una función específica, a saber, un arco bajo una estructura de adintelamiento, la pechina es su análogo en el caso de una bóveda. Ese espacio, subproducto de una adaptación arquitectónica que cumple una función, es después aprovechado secundariamente para generar una estructura de cerramiento que sirve a menudo para ornamentar el conjunto. Así pues, “el espacio surge como un subproducto necesario de la bóveda de abanico; su uso asignado es un efecto secundario” (Gould y Lewontin, 1979: 583).

Establecen así una analogía en biología evolutiva, donde un subproducto resultante de unas estructuras con función específica (adaptaciones) puede ser cooptado por ser aptativo, convirtiéndose en exaptación.

Al reconfigurar así la terminología varios problemas quedan resueltos. En primer lugar, podemos hablar de procesos de aptación, en lugar de referirlos con poco rigor como procesos de adaptación, distinguiendo dos modos de aptación, la exaptación y la adaptación, en función de si su origen se da de forma colateral o con función específica respectivamente. Queda también reconocida la existencia de subproductos que no son aptativos, si bien eso no significa que sean perjudiciales para el éxito evolutivo.

En segundo lugar, se resuelve el problema de la *preadaptación*, término que en aras de un uso riguroso sustituiremos por el de *preaptación*, el cual designa los casos de exaptación que producen adaptaciones secundarias (Gould y Vrba, 1982: 11). Dado que existen etapas incipientes de estructuras útiles surge la pregunta de por qué aparecen estas antes de especificarse su utilidad. La respuesta es que características surgidas por adaptación o como subproducto pueden establecerse por exaptación como una estructura útil. Existía, por tanto, un término (*preadaptación*) que hacía referencia a ciertos casos de exaptación, aunque no los explicaba debidamente. Por último se acuña un término para las exaptaciones que parten de subproductos, *enjutas* o *pechinas* (Gould y Vrba, 1982: 12).

Proceso		Tipo de característica		Uso	Observaciones
<p>Aptación: proceso evolutivo mediante el cual se aumenta la aptitud.</p>	<p>Adaptación: proceso mediante el cual la selección natural forma específicamente una característica para su función actual.</p>	<p>Adaptación: una característica es considerada una adaptación únicamente si ha sido desarrollada por la selección natural para desempeñar la función que actualmente desempeña.</p>		<p>Aptación: una característica es considerada una aptación si supone una mejora de la aptitud, independientemente de su origen filogenético. El estudio de este permitirá especificar si la aptación se produce mediante adaptación o mediante exaptación.</p>	<p>Función</p> <ul style="list-style-type: none"> · La adaptación puede ser cooptada si resulta aptativa para otro uso, produciendo además una exaptación. · La adaptación puede generar subproductos, estos aunque no son necesariamente aptativos pueden serlo, generando así una exaptación.
	<p>Cooptación: proceso mediante el cual una característica que resulta aptativa de forma colateral es seleccionada para un uso distinto al que desempeña por adaptación.</p>	<p>Cooptación preaptativa: una característica que surge para una función concreta (adaptación), es cooptada para un nuevo uso para el que también resulta útil.</p>	<p>Preaptación es la fase incipiente de posibles adaptaciones, surgidas como exaptaciones, y desarrolladas como adaptaciones secundarias.</p>		<p>Exaptación: una característica es considerada una exaptación cuando aumenta la aptitud del individuo, sin que esa característica haya sido promovida por la selección natural para el fin que ahora desempeña.</p>
<p>Generación de estructura no aptativa como subproducto, derivada de una adaptación.</p>		<p>Subproducto no aptativo</p>		<p>Sin aptación</p>	<p>Efecto</p> <ul style="list-style-type: none"> · El hecho de que no sea aptativo no significa que suponga un perjuicio para el éxito evolutivo, se dan varias posibilidades: que sea neutro; que perjuicios y beneficios se equilibren; que los perjuicios sean marginales.

Tabla 2. Terminología comentada de factores evolutivos relativos a la aptitud.

Cinco hipótesis posibles sobre cómo surgió lo moral en términos de selección natural.

Dependiendo de si se considera que la agencia moral aparece como adaptación, exaptación o subproducto, y sus diferentes modos, pueden distinguirse cinco hipótesis. A lo largo del presente apartado las expondremos y analizaremos su viabilidad.

Joyce señala dos corrientes de pensadores respecto a esta cuestión, haciendo una división que considero confusa y que cabe aclarar. A partir de su propuesta desarrollaremos otra más detallada y más rigurosa. Joyce bautiza como *teóricos de la enjuta* (*spandrel theorists*) a quienes consideran que la agencia moral surgió como un subproducto de otras facultades adquiridas adaptativamente. Entre ellos menciona a Nichols, Pritz, Ayala, Machery y Mallon. El nombre que les da no es adecuado, *enjuta* o *spandrel* refiere a subproductos que aumentan la aptitud por tener una utilidad que es resultado colateral en la evolución, sin embargo podría darse el caso de que algunos de estos autores consideren que la agencia moral es un subproducto no aptativo. Así pues este grupo debería llamarse *teóricos del subproducto moral* [TSM] (*moral by-product theorists*) y sería divisible en *teóricos de la enjuta moral* [TEM] (*moral spandrel theorists*) y en *teóricos del subproducto moral no aptativo* [TSMNA] (*moral non-aptive by-product theorists*) (incluyo el término *moral* para diferenciar mi clasificación de la homónima en Joyce).

Por otro lado refiere a los *nativistas morales* (*moral nativists*), respecto a los cuales sugiero un mejor término, *adaptacionistas morales* [AM] (*moral adaptationists*), ya que plantean que la agencia moral es una adaptación, desarrollada con una función específica. Entre ellos menciona a Alexander, Irons, Krebs, Dwyer, Mikhail y a sí mismo, Joyce (Joyce, 2014: 262). Plantea además dos modos en los que la agencia moral puede ser una adaptación. Uno de ellos refiere a la posibilidad de que surgiese directamente como una *adaptación*. El otro plantea que aunque el origen de la agencia moral fuese exaptativo podría haberse desarrollado posteriormente como un subproducto cooptado, generándose una exaptación y una *adaptación secundaria*.

Si esto ha ocurrido, entonces la capacidad para el juicio moral ya no es un mero subproducto sino más bien una adaptación por derecho propio (por supuesto, uno puede todavía mantener que apareció originalmente como un subproducto, pero esto es cierto prácticamente de todo lo que cuenta como una adaptación. (Joyce, 2014: 264)

Los AM podrían dividirse, por tanto, en *adaptacionistas morales primarios* [AMP] (*moral primary adaptationists*) y *adaptacionistas morales secundarios* [ASM] (*moral secondary adaptationists*). Los primeros consideran que la agencia moral surge directamente como una adaptación, con una función específica, mejorando la aptitud. Los segundos plantean que el origen de la agencia moral puede producirse a partir de una exaptación, siendo después desarrollada como adaptación secundaria, por resultar útil. La exaptación puede partir de un segundo uso de una adaptación, o bien de un subproducto que resulta útil, generando una enjuta. Así pues, algunos ASM caerían también bajo la categoría de TSM, en la medida en que el origen inicial de esa adaptación parte de un subproducto. Pueden diferenciarse, en consecuencia, dos tipos de AMS, los *AMS no-enjuta* (*non-spandrel moral secondary adaptationists*) y los *AMS pro-enjuta* (*spandrel moral secondary adaptationists*).

Grupo	Subgrupo	Observaciones	Plausibilidad
Adaptacionistas morales [AM] (<i>Moral adaptationists</i>)	Adaptacionistas morales primarios [AMP] (<i>Moral primary adaptationists</i>)	Consideran que la agencia moral surge directamente como una adaptación, que desempeñaría una función específica que aumenta la aptitud.	No es plausible (véase la justificación más abajo)
	Adaptacionistas morales secundarios [AMS] (<i>Moral secondary adaptationists</i>)	AMS no-enjuta (Non-spandrel MSA) Consideran que la agencia moral tiene origen en una exaptación, desde una adaptación primaria, sin que haya subproducto, generándose una adaptación secundaria.	
Teóricos del subproducto moral [TSM] (<i>Moral by-product theorists</i>)	AMS no-enjuta (Non-spandrel MSA)	Algunos AMS son también teóricos del subproducto moral, considerando que el subproducto que dio origen a la agencia moral se desarrolló más tarde como adaptación secundaria, por resultar aptativo.	Es plausible (véase la justificación más abajo)
	Teóricos de la enjuta moral [TEM] (<i>Moral primary adaptationists</i>)	La agencia moral surge como subproducto y resulta tener valor aptativo, sin desarrollarse posteriormente como una adaptación.	
	Teóricos del subproducto moral no aptativo [TSMNA] (<i>Moral non-aptive by-product theorists</i>)	La agencia moral surge como subproducto pero no tiene valor aptativo.	

Tabla 3. Clasificación de posturas respecto a la aparición y desarrollo evolutivo de la agencia moral. En negrita las cinco posturas posibles, que se pueden encuadrar en dos categorías generales (grupo), existiendo una subcategoría general para los dos tipos de AMS.

No podemos situar a todos los autores de forma clara en alguna de estas subcategorías. Resulta fácil clasificarlos como *adaptacionistas morales* (AM) o *teóricos del subproducto moral* (TSM), en función de si consideran que la agencia moral surge como adaptación o como subproducto. Pero muchos de ellos no se pronuncian acerca del nivel de aptitud de la agencia moral y su posible desarrollo como adaptación secundaria. Algunos casos que sí permiten esta clasificación son Ayala, que a la luz de su propuesta caería bajo la categoría de TSMNA, y Joyce, que sería AMS, si bien no ofrece datos suficientes para situarlo entre los *AMS no-enjuta* o los *pro-enjuta*.

Es preciso analizar estas posturas a la luz de todo lo expuesto hasta ahora, de este modo podremos delimitar cuáles de las explicaciones en torno al origen evolutivo de la agencia moral son plausibles (justificando así las valoraciones de la columna derecha de la Tabla 3).

a) *Adaptacionistas morales primarios (AMP)*: consideran que la agencia moral es una adaptación propiamente dicha. A la luz de todo lo expuesto a lo largo del texto podemos descartar esta opción, cuya tesis no suscriben muchos especialistas. La moralidad, más bien, aparece como resultado de la concurrencia de varias facultades desarrolladas para cumplir otras funciones.

b) *Adaptacionistas morales secundarios no-enjuta (AMS no-enjuta)*: ciertas facultades, surgidas como adaptaciones, permiten desempeñar la operación de la agencia moral, siendo cooptadas por resultar aptativa esta facultad, la agencia moral, produciéndose así una exaptación. Hemos de descartar esta opción, la agencia moral no deriva del mero uso de facultades ya existentes, sino que la concurrencia de estas forma una estructura nueva que posibilita la agencia moral. Puesto que la estructura es nueva ha de considerarse que el origen de la agencia moral se produce en un subproducto. En virtud de ello es necesario recurrir a una *teoría del subproducto moral (moral by-product theory)*.

d) *Adaptacionistas morales secundarios pro-enjuta (AMS pro-enjuta)*: la concurrencia de una serie de facultades genera como subproducto una nueva estructura que permite la agencia moral, dándose el caso de que esta además es aptativa, este subproducto es cooptado para desempeñar esa tarea, produciéndose una exaptación en la forma de *enjuta (spandrel)*. Esta exaptación se desarrolla

posteriormente como una adaptación secundaria. No descartamos esta posibilidad, resulta plausible, aunque no se ha verificado empíricamente.

e) *Teóricos de la enjuta moral (TEM)*: la concurrencia de una serie de facultades genera como subproducto una nueva estructura que permite la agencia moral, dándose el caso de que esta además es aptativa. Este subproducto es cooptado para desempeñar esa tarea, y se produce una exaptación en la forma de enjuta. Pero esta exaptación no se desarrolla posteriormente como una adaptación secundaria, permanece como un subproducto útil, es decir, una enjuta o *spandrel*. También esta posibilidad resulta plausible, pero no cuenta con apoyo empírico que permita defenderla adecuadamente.

f) *Teóricos del subproducto moral no aptativo (TSMNA)*: La concurrencia de una serie de facultades genera como subproducto una estructura nueva que permite la agencia moral, pero la agencia moral no resulta ser aptativa –sin perjuicio en términos de éxito evolutivo–, por lo tanto no se produce cooptación ni exaptación, es un mero subproducto, no aptativo. Tampoco descartamos esta posibilidad, es plausible, pero tampoco ha podido ser verificada empíricamente.

A favor de la agencia moral como subproducto evolutivo.

La facultad de la agencia moral se produce gracias al ejercicio conjunto de otras facultades y disposiciones psicológicas. Tal como hemos expuesto en el epígrafe anterior. Si bien estas existen y operan interrelacionadamente, vamos a dividir las en diferentes estratos para su exposición. Uno de estos estratos es de corte emocional, compartido con otras especies en gran medida. En él encontramos la prosocialidad (Darwin, Waal, Tomasello), dentro de la cual cabe identificar, por un lado, cierto tipo de emociones, como la culpa, la vergüenza, la gratitud, la simpatía, la estima, etc., que están presentes en humanos y encuentran correlatos en otras especies (Eisenberg, 2000; Tagney, 2007; Prinz, 2011), y, por otro, facultades prosociales, como la teoría de la mente y la empatía (Meltzoff, 2002; Torralva y Manes, 2014; Iaconi, 2009; Rizzolatti y Sinigaglia, 2006)¹⁰. Estos elementos dirigen el ejercicio de la agencia moral hacia los otros y no hacia

¹⁰ La teoría de la mente (ToM, por sus siglas en inglés —Theory of Mind—) es la capacidad de atribuir estados mentales (creencias, intenciones, deseos, conocimiento, etc.) a otros individuos y a uno mismo. Si bien está presente en humanos, salvo en algunos casos de déficits cognitivos (común en daño cerebral, espectro autista, esquizofrenia, etc.) no se ha confirmado su existencia en otras especies. Algunos estudios

uno mismo. Ciertos casos atípicos, como el de la psicopatía, muestran que la racionalidad privada de una eficiente prosocialidad no produce un agente moral en el pleno sentido de la palabra (Darwin, 1979; Haidt, 2001: 816; Prinz, 2011; Damasio; Greene; Moll; Raine; Blair,). Asimismo, el componente emocional de este estrato habilita al individuo como evaluador de conductas y situaciones, pudiendo sentir agrado o desagrado respecto a estas (Prinz). Encontramos un segundo estrato, el que permite la metacognición y los procesos de razonamiento lingüístico. Este aparece de forma tardía en la evolución, reconfigurando todos los elementos del nivel anterior. Si bien existen ciertas continuidades con otras especies en algunas de las facultades que posibilitan este estrato, el nivel de complejidad que estas alcanzan en nuestra especie y el ejercicio conjunto de las mismas hacen del nivel metacognitivo, racional y lingüístico algo exclusivo en nuestra especie. En este estrato encontramos principalmente la razón, el lenguaje y el pensamiento abstracto, aunque también están presentes facultades como la memoria y la imaginación. Este nivel permite que las evaluaciones puedan producir juicios morales respaldados racionalmente en términos de justificación.

Sostengo que la agencia moral no surgió como una adaptación propiamente dicha, para desempeñar una función concreta. Sino que la concurrencia de una serie de facultades y disposiciones biológicas que habían aparecido y se habían desarrollado para desempeñar otra función, una vez alcanzaron cierto nivel de complejidad, y sumadas a la prosocialidad, favorecieron la aparición de una nueva estructura, que se produce como efecto colateral. Esta estructura es la agencia moral, que surgió, por tanto, como un subproducto evolutivo. Esta postura me sitúa entre los TSM. No me pronunciaré acerca de si resulta aptativa o no, por ello no podré situarme en ninguna de las subcategorías correspondientes, a saber, TSMNA, TEM y AMS pro-enjuta.

Entre los TSM encontramos a Ayala, Damasio, Nichols y Prinz, entre otros. El propio Darwin parece posicionarse también a favor de la tesis que defienden estos autores. A continuación veremos cuáles son los principales argumentos que me han llevado a posicionarme a favor de esta postura.

Para comprender el origen de la agencia moral como subproducto es preciso atender al hecho de que el cerebro humano es una estructura complejísima, en función de

sugieren que hay evidencia suficiente para apoyar la tesis de que algunas especies de animales no humanos sí presentan esta capacidad, especialmente en algunas especies de primates, cetáceos, córvidos y canes.

lo cual resulta propenso a la aparición de subproductos evolutivos y exaptaciones. Gould y Vrba afirman lo siguiente:

[...] aunque sin duda construido por la selección para algún complejo conjunto de funciones, [*el cerebro*]¹¹ puede, como un resultado de su intrincada estructura, trabajar en un número ilimitado de formas muy poco relacionadas con la presión selectiva que lo construyó. Muchas de estas formas pueden llegar a ser importantes, si no indispensables, para la futura supervivencia en contextos sociales posteriores [...]. Sin embargo la utilidad actual no acarrea implicaciones automáticas sobre el origen histórico. La mayoría de lo que el cerebro hace ahora para mejorar nuestra supervivencia radica en el dominio de la exaptación –y no nos permite hacer hipótesis sobre los senderos selectivos de la historia humana. (Gould y Vrba, 1982: 13)

La complejidad cerebral permite a este órgano trabajar en infinidad de vías para las que no fue construido por la selección natural, favoreciendo la posibilidad de subproductos y exaptaciones. La posibilidad de que la agencia moral sea un subproducto, no necesariamente aptativo, de facultades que fueron adquiridas y desarrolladas para desempeñar otras funciones, está ya latente en los textos del propio Darwin. Joyce respalda esta idea en un artículo donde aborda esta cuestión.

Darwin ve el sentido moral como emergente (inevitablemente) a partir de otros rasgos: “instintos sociales” combinados con “facultades intelectuales”. Estas últimas facultades que continúa mencionando son la memoria, el lenguaje y el hábito. Esto plantea la posibilidad de que Darwin no ve el sentido moral como una discreta adaptación psicológica sino más bien como un subproducto de otros rasgos evolucionados. De hecho, parece sabiamente mantenerse alejado de pronunciarse en este asunto. Cuando se centró en los instintos sociales en general (más que en el sentido moral en particular), escribe que “es... imposible decidir en muchos casos si ciertos instintos sociales han sido adquiridos a través de la selección natural, o son el resultado indirecto de otros instintos y facultades”. (Joyce, 2014: 262)

Atendiendo a los textos del propio Darwin puede respaldarse la postura de Joyce. Darwin considera que la prosocialidad sumada al ejercicio de ciertas facultades, habiendo alcanzado estas determinado grado de complejidad, harían que cualquier animal adquiriese el sentido moral del que habla en sus textos, que podemos identificar con la facultad de la agencia moral. Esta es precisamente la idea que defiende Ayala, quien considera que las habilidades intelectuales de nuestra especie, habiendo alcanzado cierto umbral de complejidad, permiten, dada su concurrencia, la aparición del comportamiento

¹¹ Los corchetes son míos, en aras de especificar cuál es el sujeto de la oración, ya que queda omitido al seleccionar la cita.

ético, el cual es un resultado fortuito que no surge por ser aptativo, sino por la concurrencia de esas otras facultades que sí aparecieron para desempeñar ciertas funciones. En la primera parte del epígrafe anterior hemos expuesto de forma amplia la postura tanto de Darwin como de Ayala respecto a esta cuestión.

Damasio comparte la idea de la agencia moral surgida como subproducto de otras facultades. *En busca de Spinoza* contiene un capítulo sobre “Neurobiología y comportamientos éticos” en el que Damasio expone que no hay un centro moral en el cerebro, simplemente, porque la moralidad no apareció para desempeñar una función específica, sino como resultado del ejercicio de otras facultades concurrentes. (Damasio, 2005: 160)

También Greene y Haidt (2002) y Greene, de forma transversal en toda su propuesta sobre el proceso dual del juicio moral, defienden la idea de que no existe un centro moral en el cerebro. Los experimentos neurocientíficos de Greene en sus investigaciones revelan la participación de diferentes áreas del cerebro, vinculadas a las emociones y a la razón en distinto grado, según el tipo de dilemas a los que se someta el sujeto de experimentación. Lo cual revela que diversas áreas cerebrales desarrolladas para funciones distintas al dominio moral son ejercitadas de forma conjunta a la hora de producir el juicio moral.

Respecto a esa concurrencia y sus efectos es muy interesante la aportación de Nichols, quien —tal y como nos muestra Joyce— plantea que la agencia moral es posible gracias a la capacidad de usar imperativos no-hipotéticos y un mecanismo afectivo y emocional para responder al sufrimiento de otros, la prosocialidad. Estos elementos pueden ser adaptaciones, pero gestan como subproducto (*inter-alia*) la agencia moral.

Sin embargo, es claramente menos plausible que la capacidad para los juicios morales básicos sea una adaptación en sí. Es más probable que los juicios morales básicos sean una especie de subproducto de (*inter-alia*) los mecanismos innatos afectivos y los mecanismos innatos de comprensión de reglas. (Joyce, 2014: 263)

Asimismo, Prinz sostiene que no existe un dispositivo biológico para la moralidad, esta es resultado de la interacción de otros dispositivos o facultades, que surgieron para desempeñar funciones distintas. La propuesta de Prinz falla al indicar qué dispositivos son esos que permiten la aparición de lo moral, señala solo las emociones, la imitación y

el pensamiento abstracto, olvidando otros muchos elementos a los que hemos referido en apartados anteriores. Si bien su propuesta es deficitaria en esa cuestión, su diagnóstico sobre el origen de la agencia moral es, desde nuestra postura, acertado (Prinz, 2011; Sripada y Stich, 2005).

La adquisición de reglas morales puede no depender de ninguna especie de dispositivo de adquisición de la moralidad (Sripada & Stich, 2005), pero en cambio pueden derivar de recursos cognitivos que evolucionaron para otros propósitos [...]. Así como la religión puede surgir en todas las culturas sin un módulo de la religión, la moralidad puede ser un subproducto de capacidades que no son específicas del dominio moral. (Prinz, 2011: 11)

Prinz coincide con Damasio en que no existe un área cerebral destinada específicamente a la moralidad, lo cual es un indicativo de que lo moral es posible gracias a la interacción de diferentes áreas con diferentes funciones, surgidas por motivos distintos a la agencia moral y que posibilitaron esta más tarde, dada su concurrencia y alcanzado cierto nivel de complejidad.

Tomando como plausibles las tres posturas que engloba la categoría de TSM (teóricos del subproducto moral) habría que resolver si la agencia moral es un subproducto aptativo o no aptativo. De serlo, habría que resolver también si fue desarrollado posteriormente como adaptación secundaria. Lamentablemente no podemos ofrecer una respuesta fiable. Es necesaria una investigación en la que ciencia y filosofía profundicen en la cuestión para perfilar los escenarios posibles en los cuales la moralidad pudo surgir, intentando ver en qué medida pudo mejorar el nivel de aptación en nuestros ancestros.

A continuación, consideraremos varios argumentos que ponen en cuestión que la agencia moral sea aptativa, exponiendo también algunas de las objeciones posibles a tales argumentos:

1) *El nivel de aptación depende de factores contingentes*: que la aptitud mejore depende de los diferentes contenidos morales, que son contingentes, por tanto la agencia moral no es aptativa *per se*. *Objeción*: por lo general, aunque los contenidos morales difieran entre culturas, suelen favorecer la vida grupal y la prosocialidad, esta tendencia es consecuencia de ciertas predisposiciones biológicas subyacentes en nuestra conducta.

2) *Puede ser perjudicial*: hay casos en que la agencia moral puede perjudicar o destruir el individuo, el grupo o la especie, por convicciones y prácticas morales.

Objeciones: a) este argumento es insuficiente por sí mismo, el hecho de que una facultad o característica tenga usos que atenten contra la supervivencia y la reproducción no significa que no sea aptativa, hay adaptaciones (ej. la mano) que presentan el mismo riesgo; b) la tendencia a la prosocialidad haría que los casos en que la moralidad resulte perjudicial sean marginales y anecdóticos.

3) *Debe ofrecer una ventaja selectiva*: para ser aptativa la agencia moral tiene que ofrecer una ventaja evolutiva al agente moral. Es preciso que prescindir de ella no resulte posible dado cierto escenario de selección natural. Debe funcionar más eficientemente que los mecanismos ya existentes que regulan la vida social en la especie que deviene moral, si es igual en eficiencia, no sería imprescindible.

Objeción: una vez aparece la agencia moral el escenario evolutivo en que el sujeto se encuentra se reconfigura, surgiendo problemas nuevos, cuya resolución precisaría la existencia de la agencia moral, que sí sería aptativa en este caso.

Por consiguiente, a pesar de posicionarnos entre los TSM, considerando que la agencia moral tiene origen como subproducto evolutivo, no tenemos evidencias claras sobre su carácter aptativo o adaptativo, en caso de tenerlo. Corresponderá a ulteriores investigaciones profundizar en el problema para arrojar respuestas más claras respecto al mismo.

La agencia moral en perspectiva naturalista y sus implicaciones para la ética.

Una vez tratadas estas cuestiones en relación a la agencia moral hemos de preguntarnos qué implicaciones tienen para la ética. En primer lugar hay que aclarar que el hecho de que la agencia moral surja como un subproducto no desprestigia la relevancia de lo ético. La categoría biológica de subproducto (o *coproducto*) no presenta ninguna connotación peyorativa. En segundo lugar hay que aclarar que si la agencia moral no fuese aptativa tampoco quedaría menoscabada la importancia de lo moral para el ser humano. La ética, en tanto que proceso metacognitivo que busca la justificación de la acción, tiene una estructura de sentido independiente a las bases biológicas que la posibilitan. Una vez el ser humano se constituye como agente moral se abre un espacio de reflexión en torno a cuestiones como el deber, la obligación, el bien y la justicia. Así

pues, la ética, posibilitada por la agencia moral en tanto que facultad natural, se constituye como una operación cuyo sentido dista de tal naturaleza, aunque su ejercicio dependa de ella (Cortina, 2007, 2012a; Habermas, 2006; Korsgaard, 2000). Una facultad que aparece biológicamente como subproducto pasa, así, a configurar uno de los ejes de las sociedades humanas en virtud de la relevancia que cobra lo ético para ese animal que ha devenido agente moral.

La naturaleza de la agencia moral y el sentido de la ética presentan, por tanto, una distancia que imposibilita la reducción de una de las esferas a la otra. Mientras que el primer caso corresponde al campo de la ciencia, que versa sobre cómo es el mundo, el segundo corresponde al campo de la filosofía moral o ética, que versa sobre cómo debe ser, tratando de resolver el problema metacognitivo de justificación de la acción. La naturaleza no es intrínsecamente normativa, es por ello que se hace preciso recurrir a la ética para resolver el problema de la fundamentación, en el que la ciencia es infecunda.

Capítulo 3. Sobre la naturalización del juicio moral

La investigación neurocientífica de lo moral presenta como uno de sus elementos centrales el juicio moral. Un vasto número de las investigaciones en neuroética orbitan en torno a aquello que sucede en el cerebro cuando un sujeto produce y emite un juicio moral. No obstante, merece la pena adoptar una actitud crítica en la medida en que algunos resultados de estas investigaciones pretendan otorgar a la ciencia más competencias de las que le corresponden, algo que se produce, por ejemplo, cuando se trata de extraer conclusiones normativas a partir de experimentos científicos. Muchas son las preguntas que surgen. ¿Qué es eso a lo que llamamos juicio moral? ¿Cómo funciona? ¿Puede una investigación de corte naturalista dar cuenta de cuestiones fundamentales de la ética mediante el estudio del juicio moral a nivel cerebral? ¿Qué tipo de experimentos se han llevado a cabo? ¿Qué tipo de conclusiones procede extraer de tales experimentos? ¿Basta la investigación de los juicios morales para extraer conclusiones representativas sobre qué es la moralidad y cómo funciona?

En el presente capítulo abordaremos estas preguntas en dos partes. En la primera parte expondremos algunas de las propuestas más representativas en torno a la investigación del juicio moral, centrándonos en los logros científicos que se han podido cosechar y en los principales obstáculos encontrados. Atenderemos a las propuestas de Haidt, Greene, Damasio, Mikhail y Cortina. En la segunda parte retomaremos el caso de Joshua Greene, para someter a crítica algunas de sus conclusiones filosóficas. Sus investigaciones neurocientíficas en psicología moral componen un caso paradigmático en torno a la investigación naturalista del juicio moral, cometiendo además algunos de los errores reduccionistas que venimos denunciando. Greene no es el único que comete este tipo de errores, pero merece una especial atención ya que, en primer lugar, es un caso muy representativo de la naturalización del juicio moral y de los riesgos que esta entraña, y, en segundo lugar, sus investigaciones han recibido un reconocimiento y un prestigio exacerbado. Atender a él nos da la oportunidad de ejercer una crítica que es muy necesaria sobre su propuesta.

3.1. El juicio moral desde la perspectiva científica

Teorías del proceso dual.

El modelo intuicionista social de Haidt.

Uno de los textos seminales en neuroética en torno al juicio moral corresponde a Jonathan Haidt, es un artículo publicado en 2001 cuyo título es “The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment”, cuya propuesta es ampliada por el propio autor en *The righteous mind*. La tesis principal de Haidt es que por lo general el juicio moral se produce de forma intuitiva, mientras que la razón interviene solo posteriormente. El razonamiento moral sería, pues, no la causa del juicio, sino una construcción *post hoc*. Con ello Haidt esgrime una crítica a los enfoques éticos racionalistas que tradicionalmente han predominado en la filosofía moral. Según estos enfoques el juicio y el conocimiento moral son resultado de procesos de razonamiento y reflexión. La propuesta de Haidt desafía estos modelos, ya que relega el papel de la razón, en la mayoría de los casos, a tareas de racionalización, no de fundamentación racional, por ejemplo buscando apoyo racional para aquellos juicios hacia los cuales nos sentimos intuitivamente inclinados.

La afirmación central del modelo intuicionista social es que el juicio moral es causado por rápidas intuiciones morales y es seguido (cuando es necesario) por ex post facto razonamientos morales lentos. (Haidt, 2001: 817)

Tres son los elementos centrales de su tesis. Ocupando un papel central, encontramos el juicio moral. En relación a cómo este se produce encontramos el razonamiento moral y la intuición moral. Se hace preciso atender a la definición que de estos ofrece Haidt:

- *Juicio moral*: el juicio moral es una evaluación de las acciones o el carácter de una persona, esta evaluación se produce siempre en el seno de una cultura y en relación a una serie de virtudes -o en sentido más laxo, una serie de contenidos morales- tenidos por obligatorios en el seno de esa cultura o subcultura.
- *Razonamiento moral*: el razonamiento moral, según las propuestas racionalistas de la moral, es una actividad mental consciente que consiste en transformar cierta información —sobre otras personas, sobre sus acciones, en relación a nuestros

valores, etc.— con el objetivo de producir un juicio moral. El hecho de que sea un proceso consciente implica que es intencional, con esfuerzo y controlable. En el proceso quedaría excluida cualquier forma de intuición. El razonador moral operaría en varios pasos que darían como resultado el juicio moral. En este sentido el razonador se comportaría como un científico. Este será el enfoque del razonamiento moral que Haidt pondrá en cuestión.

- *Intuición moral*: Cualquier tipo de intuición se produce de forma rápida, automática y sin esfuerzo, siendo el resultado accesible a la conciencia, pero no el proceso que lo ha producido. La intuición moral, por tanto, puede definirse como la repentina aparición de un juicio moral en la conciencia, sin presentar procesos conscientes de búsqueda y balance de evidencias, o de inferencia de conclusiones.

Llegados a este punto estamos en posición de comprender la propuesta de Haidt, a la que este bautiza como *modelo intuicionista social* (*social intuitionist model*). Este se divide en 6 eslabones, siendo los 4 primeros aquellos que componen el núcleo de la propuesta:

1. *El eslabón del juicio intuitivo*. El modelo propone que el juicio moral aparece en la conciencia automáticamente y sin esfuerzo como resultado de intuiciones morales. [...]
2. *El eslabón del razonamiento post hoc*. El modelo propone que el razonamiento moral es un proceso esforzado, realizado después de que un juicio moral sea hecho, en el cual una persona busca argumentos que apoyarán un juicio ya hecho [...].
3. *El eslabón de la persuasión razonada*. El modelo propone que el razonamiento moral es producido y enviado verbalmente para justificar el juicio moral ya hecho de uno ante los otros. Tal razonamiento a veces puede afectar a otras personas, aunque las discusiones y argumentos morales son infames por la rareza con que tiene lugar la persuasión. Debido a que las posiciones morales siempre tienen un componente afectivo para ellos, se hipotetiza que la persuasión razonada no funciona por proporcionar argumentos lógicamente convincentes sino al activar nuevas intuiciones con valencia afectiva en el oyente. [...]
4. *El eslabón de la persuasión social*. Debido a que las personas están muy en sintonía con el surgimiento de normas grupales, el modelo propone que el mero hecho de que amigos, aliados y conocidos hayan emitido un juicio moral ejerce una influencia directa sobre otros, incluso si no se utiliza una persuasión razonada. [...]
5. *El eslabón del juicio razonado*. Las personas a veces pueden razonar el camino hacia un juicio por la pura fuerza de la lógica, anulando su intuición inicial. En tales casos, el razonamiento es realmente causal y no puede decirse que sea el "esclavo de las pasiones". Sin embargo, se

presume que tal razonamiento es raro, ocurriendo principalmente en casos en los que la intuición inicial es débil y la capacidad de procesamiento es alta. [...]

6. *El eslabón de la reflexión privada*. En el momento de pensar sobre una situación una persona puede espontáneamente activar una nueva intuición que contradice el juicio intuitivo inicial. [...] (Haidt, 2001: 818-819)

Estos seis puntos ilustran las ideas básicas que Haidt defiende, para una mejor comprensión del modelo será preciso comentar cada uno de estos puntos detenidamente. El primer punto o eslabón del modelo refiere al juicio intuitivo. Según Haidt la mayoría de los juicios morales que producimos son de este tipo, aparecen fortuitamente en la conciencia como resultado de procesos de los que no somos conscientes. Estos procesos vienen condicionados por predisposiciones tanto biológicas como culturales que hacen que experimentemos aprobación o desaprobación moral hacia ciertos hechos, acciones o caracteres.

El segundo de los eslabones corresponde al papel que la razón ejerce *post hoc*, según Haidt la función que la razón desempeña generalmente consiste en un proceso de búsqueda de argumentos que permitan respaldar un juicio que ya viene dado por la intuición. Posteriormente en el texto Haidt dirá que la razón funciona, por tanto, más como un abogado que como un científico, la metáfora es muy ilustrativa. El razonamiento moral no consistiría en un proceso lógico honesto y riguroso que buscase el juicio más adecuado, sino en la articulación de una serie de argumentos cuyas premisas desemboquen en una conclusión que encarne y respalde el juicio moral que ya tenemos. En ocasiones sucede que los razonadores morales no pueden dar cuenta racional de sus juicios, se produce entonces el llamado desconcierto moral. Por sorprendente que parezca, en estos casos el juicio no es desechado o puesto en tela de juicio, sino que, manteniendo la confianza en la intuición moral, el agente sostiene posturas tales como que aun no sabiendo por qué, sabe que algo está bien o mal, según el caso.

Hasta aquí hemos podido ver el componente intuicionista del modelo, pero este también es social, según expone su autor. Los eslabones tres y cuatro hacen referencia a esta dimensión social. Cuando el juicio se produce y se mantiene de forma privada la necesidad de racionalizar su validez no siempre se presenta. Es en la esfera social donde se nos exige y donde exigimos justificar por qué algo es tomado como bueno o malo, como justo o injusto. Así pues, la búsqueda de argumentos que apoyen el juicio no responde solo a un interés privado por respaldar nuestras convicciones morales, sino que

responde a la necesidad social de persuadir a otros bien de que nuestros juicios cuentan con razones para mantenerlos, bien de que los suyos no lo hacen, o bien de ambas cosas. A menudo estos procesos sociales están interrelacionados, una justificación exitosa de un juicio puede ser el desencadenante de que en otro u otros individuos se produzca una nueva intuición moral, es frecuente que al comulgar con las convicciones que un interlocutor nos presenta tengamos que renunciar a algunas propias con las cuales se contradicen aquellas. También en este caso el papel de la intuición es tanto o más importante que el de la razón, si bien un argumento bien construido debería hacer cambiar la opinión de los otros, no siempre sucede así, y aun reconociendo la validez del argumento se mantienen ciertas convicciones que se contradicen con él. No todos los interlocutores se someten a la fuerza de la razón. Los elementos afectivos y emocionales parecen tener más éxito, la persuasión suele operar con ellos a menudo, tratando de provocar una intuición en el otro, que no necesariamente precisa estar respaldada argumentativamente.¹²

Una vez expuestos los rasgos principales del modelo, Haidt aporta cuatro razones por las que debemos dudar de la importancia causal de la razón en el juicio moral. Dado que el autor está poniendo en tela de juicio los modelos racionalistas del juicio moral, estas cuatro razones suponen puntos centrales de su crítica. Si bien ya han sido planteadas en algunos de los seis puntos que hemos mencionado y que componen el modelo, cabe atender a cada una de ellas más detenidamente.

La primera de las razones tiene que ver con el llamado problema del proceso dual. Las teorías del proceso dual plantean que existen dos sistemas de procesado que operan cuando una persona hace un juicio moral o trata de resolver un problema, estos son el proceso racional y el proceso intuitivo. Si bien el primero de ellos ha recibido históricamente más atención, en las últimas décadas se está reconociendo el papel fundamental del proceso intuitivo. Este tiene preminencia en diversos aspectos, pues el sistema afectivo, sobre el que se asienta el proceso intuitivo, apareció primero

¹² La ética de la razón cordial ofrece una propuesta de filosofía moral que permite compaginar los elementos racionales con los emocionales propios de los razonadores morales en el seno de una comunidad de comunicación, permitiendo resolver los principales problemas que Haidt denuncia. Nuestra investigación, atendiendo a las aportaciones de las neurociencias y a la propuesta filosófica de una ética de la razón cordial (Cortina, 2009), ofrecerá las claves metodológicas para una neuroética crítica en clave cordial.

filogenéticamente, aparece primero ontogenéticamente¹³ y opera de forma más rápida y automática a la hora de evaluar o emitir un juicio. Además es más fuerte e irrevocable cuando los dos sistemas entran en conflicto. El hecho de que desde un contexto social se fuerce al agente moral a buscar razones que apoyen su juicio puede dar a este el aspecto de que es un juicio racional, pero de hecho es un juicio intuitivo que ha sido racionalizado *post hoc*.

En segundo lugar, Haidt plantea el problema de la motivación de la razón, sugiriendo que el proceso racional se parece más a un abogado que defiende a un cliente que a un juez o un científico en busca de la verdad. Haidt asegura que el eslabón de razonamiento *post hoc* (número 2) es más importante que los de juicio razonado y reflexión privada (números 5 y 6). Además, analiza los dos motivos principales para que la razón intervenga *post hoc*, uno de ellos es por la relación con otras personas, que suele implicar una exigencia de razonamiento en la defensa de nuestros juicios ante ellas, el segundo es por motivos de coherencia, si se da el caso de que un juicio intuitivo se contradice con otros juicios o creencias propias. Además, Haidt señala que es muy frecuente buscar indicios que apoyen nuestros juicios, desestimando aquellos que los contradicen, así como hacer las preguntas que propician una respuesta que apoyaría el juicio o creencia que ya se tiene intuitivamente. Por todo ello Haidt plantea la razón como un abogado en defensa de un cliente, más que como alguien que busca la verdad.

En tercer lugar, Haidt señala que las justificaciones *post hoc* del juicio intuitivo realizadas por la razón pueden generar la ilusión de razonamiento objetivo. Plantea que, siendo esto así, dos ilusiones nos gobiernan. La primera de ellas la bautiza como *wag-the-dog-tail illusion*, esta consiste en pensar que nuestro juicio moral (el perro), es movido por la razón (la cola), cuando en realidad la razón interviene solo posteriormente, como la cola del perro, que se mueve solo cuando este ya se siente feliz, y no al revés. De igual modo la razón se mueve cuando ya existe un juicio, y no al revés, al menos en la mayoría de los casos. La segunda ilusión es la llamada *wag-the-others-dog's-tail illusion*, y consiste en pensar que siempre tendremos éxito en convencer a otra persona de que su juicio es erróneo tan solo atendiendo a sus razones y tratando de cambiarlas, esto, dice Haidt, es como pensar que mover la cola del perro va a hacerle sentirse feliz. Para cambiar

¹³ Frans de Waal y Tomasello, y de forma más sumaria Darwin y Ayala, así como Damasio, Churchland y otros, ofrecen apoyo a esta tesis en relación a la aparición filogenética y ontogenética de los sistemas afectivo y racional.

las creencias y el juicio de los demás es mucho más adecuado intervenir también en los aspectos emocionales e intuitivos relacionados con el juicio. No obstante, Haidt admite que el modelo intuicionista social contempla que en condiciones adecuadas de juicio un razonador debería cambiar el juicio si se le exponen buenas razones, aunque no se sienta intuitivamente inclinado a ello, lamentablemente esto sucede solo en un reducido número de casos.

Por último, en cuarto lugar, Haidt señala que la acción moral guarda una relación más fuerte con las emociones que con la razón. Suelen ser estas las que mueven al sujeto a actuar. El altruismo, la empatía, la culpa, la vergüenza, etc., son factores emocionales que tienen una mayor fuerza sobre la acción que la razón, a pesar de que esta también pueda ejercer cierta influencia. El hecho de que la razón y el lenguaje sean usados para expresar y comunicar el juicio, como el movimiento de la cola de un perro, que también sirve para expresar y comunicar, no significa que los motivos internos que conducen a la acción sean racionales. Para comprender de forma profunda la moralidad humana debemos investigar los motivos emocionales de la acción, del mismo modo que para comprender la conducta de un perro no basta con atender a su cola.

Si bien Haidt continúa reflexionando sobre algunos aspectos del modelo que propone, hasta aquí hemos atendido a los rasgos principales de su propuesta y a la relación que en ellos guardan los procesos de intuición y razonamiento, que serán clave para enlazar su propuesta con la de Greene.

Greene y la teoría del proceso dual.

En 2001, el mismo año que Haidt publicó su célebre artículo, Greene publicó el primero de una serie de artículos que también planteaban la necesidad de atender a los factores emocionales al abordar la cuestión del juicio moral, titulado “An fMRI Investigation of Emotional Engagement in Moral Judgment”. A diferencia de Haidt, Greene se ocupaba de la cuestión desde una perspectiva científica y experimental que hacía uso de los últimos avances en neurociencias. Atenderemos en este apartado a estos logros científicos, dejando para más adelante el problema que suponen las conclusiones normativas que de estos extrae Greene, cuestión a la que atenderemos desde una perspectiva crítica.

El artículo parte de una serie de dilemas éticos que plantean un problema. Greene explica que ante dilemas morales similares y en los que el cálculo de resultados es aparentemente el mismo, la decisión de los agentes morales varía. Nos plantea el conocido “trolley dilemma”, o dilema del tren, y alguna de sus variantes. Este dilema plantea que un tren se dirige hacia cinco personas que se encuentran en las vías por las que este se desplaza, las cuales morirán arrolladas, algo que puede evitarse accionando una palanca y desviando el tren hacia otra vía, en la que desafortunadamente hay una persona. La respuesta de la mayoría de los sujetos del experimento, si se les pregunta qué harían, es accionar la palanca. Sin embargo, en una de las variantes del dilema, en lugar de evitar el accidente accionando una palanca, la única opción consiste en arrojar a las vías a una persona que se halla en un puente sobre las vías, cuyo cuerpo detendría el tren, muriendo. En este caso, no obstante, la respuesta mayoritaria es que no arrojarían a esta persona a las vías, aun si el resultado es la muerte de otras cinco personas. Greene se pregunta por qué varía la respuesta al variar las condiciones del dilema, aun si se mantienen los mismos resultados.

Para dar respuesta a esta pregunta Greene diseña un experimento que monitoriza la actividad neuronal de los sujetos implicados en la decisión. La hipótesis de Greene es que en la primera de las formas del dilema presenta un cálculo de los resultados más racional que la segunda formulación del dilema, en el primer caso el sujeto no se siente personalmente implicado en la decisión, mientras que en la variante del puente la decisión de arrojar a alguien a las vías hace que el sujeto se sienta implicado personalmente en la muerte del individuo. La tesis de Greene es que precisamente esa variación entre los factores racional y emocional que intervienen en el juicio hacen que el resultado de la decisión difiera.

Esta hipótesis concerniente a estos dos casos sugiere una hipótesis más general sobre el juicio moral: algunos dilemas morales (aquellos relevantemente similares al dilema del puente) involucran el procesamiento emocional en mayor medida que otros (aquellos relevantemente similares al dilema del tranvía), y estas diferencias en el compromiso emocional afecta los juicios de las personas. (Greene, 2001: 2106)

El experimento incluye una variedad de dilemas similares clasificados en tres categorías: impersonales, personales y no morales. Los resultados de la investigación muestran que efectivamente las áreas del cerebro que intervienen en los dilemas clasificados como personales y en los dilemas clasificados como impersonales varían. En

los personales toman parte áreas del cerebro donde se procesa la emoción (córtex prefrontal, córtex prefrontal ventromedial y amígdala), mientras que en los impersonales intervienen principalmente áreas asociadas al razonamiento y al cálculo (córtex prefrontal dorsolateral). Los dilemas impersonales resultan, además, muy similares a los dilemas no morales en lo que a actividad cerebral refiere.

Un año después, en 2002, ve la luz un artículo firmado por Haidt y Greene, en el cual el modelo intuicionista social del primero es complementado por los experimentos neurocientíficos del segundo. La afinidad de la propuesta de ambos da lugar a esta simbiosis entre los autores. El modelo de Haidt encuentra en su colega un apoyo que otorga un carácter científico y experimental a su propuesta. Y viceversa, Greene encuentra en Haidt un modelo del juicio moral que ofrece un sustento teórico a sus experimentos sobre el papel de la emoción en el mismo.

El artículo, titulado “How (and where) does moral judgment work?”, comienza recapitulando algunas de las propuestas racionalistas en psicología moral, donde se cita a Freud, Kohlberg y Piaget como algunos de los representantes de la revolución cognitiva que tuvo lugar en la psicología y que sufrió un especial auge en la segunda mitad del siglo pasado. El artículo pretende reivindicar la importancia de una revolución afectiva que ya viene dibujándose desde los 80, para incorporar los factores afectivos a los racionales en la psicología, y en este caso especialmente en la psicología moral. En este marco se presenta el modelo intuicionista de Haidt como uno de los exponentes más actuales de esta tendencia, ya que incorpora el factor emocional y delimita el papel y las funciones de la razón. Se citan como apoyo una serie de experimentos y autores entre los cuales se destacan las aportaciones de Greene. Entre estos autores encontramos a Antonio Damasio, cuyas investigaciones en torno al daño cerebral revelan que los déficits en el procesamiento de emociones suponen problemas en el procesamiento del juicio moral (Damasio, 1994, 2005, 2008). Otros autores que reportan datos similares son Anderson (1999), Graffman (1996) y Raine (2000). Como un precursor de los experimentos de Greene se cita a Moll (2001; 2002a; 2002b), quién lleva a cabo experimentos de neuroimagen muy similares, pero sin profundizar tanto en la cuestión moral. Tras ello se abordan con detalle los experimentos de Greene, a los que ya hemos atendido.

Los experimentos detallados en estos dos artículos (Greene, 2001; Greene y Haidt, 2002) se ven ampliados en ulteriores publicaciones de Greene. Si bien estos experimentos

son presentados inicialmente al margen de consideraciones filosóficas, poco a poco Greene irá incluyendo estas en su propuesta. De esta cuestión nos ocuparemos en el siguiente epígrafe. En lo que concierne a las aportaciones meramente científicas de Greene, las principales tesis y conclusiones que este plantea en sus investigaciones, tanto en los dos artículos comentados como en otras publicaciones posteriores, son las siguientes (la referencia que aparece es aquella en la que se presenta la idea o tesis por primera vez):

- Existen dos sistemas de procesamiento en el juicio moral, uno asociado a factores emocionales, otro asociado a factores racionales (Greene, 2001).
- La diferente respuesta de los sujetos a dilemas muy similares se debe a la variación en el nivel de intervención de los dos sistemas de procesado (Greene, 2001).
- Los dilemas en los que se ve implicado el sujeto de forma personal presentan una mayor actividad de áreas cerebrales asociadas a la emoción (Greene, 2001).
- Los dilemas de carácter impersonal en los que prevalece un cálculo de las consecuencias presentan una mayor actividad de áreas cerebrales asociadas al control cognitivo y la razón (Greene, 2001).
- Los dilemas clasificados como no morales son muy similares a los dilemas clasificados como impersonales en lo que concierne a la actividad cerebral que se asocia a ellos (Greene, 2001).
- Existen evidencias científicas que respaldan el modelo intuicionista social de Haidt, los experimentos de Greene revelan la importante actividad de las áreas cerebrales asociadas a la emoción en los procesos de juicio moral (Greene y Haidt, 2002).
- Las intuiciones morales, asociadas a la emoción, vienen condicionadas tanto por la selección natural como por factores culturales (Greene y Haidt, 2002).
- Otras investigaciones reconocidas (Damasio, Anderson, Graffman, Raine, Moll) respaldan las tesis de Greene y Haidt (Greene y Haidt, 2002).
- La razón tiene un papel importante en el procesado de dilemas impersonales, pero también en el de dilemas personales donde las consideraciones razonadas y las intuiciones morales entran en conflicto (Greene y Haidt, 2002).
- En muchos casos, no en todos, el juicio moral implica el uso de procesos específicamente dedicados a la cognición social, especialmente en el uso de teoría

- de la mente, consistente en la representación de la mente de otros (Greene y Haidt, 2002).
- Algunas emociones son más centrales en nuestra vida moral que otras (como la compasión y la culpa, entre otras), pero todas pueden intervenir en el juicio moral bajo ciertas circunstancias (Greene y Haidt, 2002).
 - No hay una parte moral en el cerebro, sino que en el juicio moral intervienen diferentes partes del cerebro que desempeñan además otras funciones no morales. Entre esas partes encontramos áreas de procesado tanto emocional como cognitivo y racional (Greene y Haidt, 2002).
 - La evolución nos predispone a procesar como personales los dilemas morales con individuos que sentimos más cercanos, mientras que los dilemas son percibidos de forma más impersonal cuando los individuos implicados parecen estar menos vinculados al sujeto (Greene, 2003).
 - La evolución ha generado mecanismos para que mediante una fenomenología perceptual tengamos intuiciones sobre lo bueno y lo malo, aunque estas no presenten un correlato natural en el mundo externo, sino que sean producto de la mente. Estas intuiciones, automáticas y emocionales, conducen a la generación de juicios morales basados en la emoción (Greene, 2003).
 - Existe una distinción real entre los procesos de emoción y cognición, pero es precisa una investigación ulterior (Greene, 2004)
 - Las intervenciones de procesos cognitivos y de racionalización en el juicio moral, incluso en casos en los que el juicio es personal, se caracterizan por funcionar como un cálculo de utilidad, que en determinados casos puede inhibir el factor emocional e imponerse (Greene, 2004).

Las tesis expuestas anteriormente componen el corpus principal de la propuesta de Greene, si bien este se ve ampliado por algunas cuestiones secundarias en diversas publicaciones, tratando temas como el funcionamiento del proceso dual cuando se produce cierto daño cerebral (Greene, 2007a), el papel que juega la intención en el proceso de juicio moral (Greene, 2009b), el papel de las representaciones de probabilidad y magnitud a la hora de hacer el juicio moral (Greene, 2010), o cómo funciona el razonamiento y la reflexión en el juicio moral en función de factores como el nivel de argumentación o el tiempo de deliberación (Greene, 2011).

En un reciente artículo de 2014, “The Cognitive Neuroscience of Moral Judgment and Decision Making”, Greene hace un recorrido por las diferentes investigaciones neurocientíficas en relación al juicio moral, con la intención de ofrecer una vista panorámica de qué puede decir la neurociencia sobre el juicio moral. Ofrece pocas conclusiones novedosas, poniendo especial atención en el hecho de que no existen áreas cerebrales específicamente morales, sino un aprovechamiento de diferentes áreas para procesar el juicio moral. Además, señala que la principal tarea de la neurociencia cognitiva consiste en atender a la relación función-estructura. En un artículo de 2015 Greene habla sobre el auge de la investigación en el campo de la cognición moral. En este artículo nombra las principales cuestiones abordadas en dicho terreno, reconociendo, además, la influencia de la filosofía. También especula sobre el futuro de la disciplina, sin aportar ninguna novedad en relación al estudio del juicio moral.

Algunos enfoques alternativos al proceso dual.

Antonio Damasio, complejidad y marcadores somáticos en el juicio moral.

Si bien Greene cita a Damasio como apoyo para su teoría del proceso dual, haciendo referencia a sus investigaciones sobre el célebre caso de Phineas Gage (Damasio, 1994, 2005, 2008), lo cierto es que las investigaciones de este autor van mucho más allá de lo que Greene hace ver. En el ámbito de la neuroética, el reconocimiento del papel de la emoción en el razonamiento y el juicio moral es un denominador común en la mayoría de autores. Lo que debemos preguntarnos es si las investigaciones de Damasio sirven de apoyo a la teoría del proceso dual tal y como lo presenta Greene, o si en realidad plantean algunos aspectos del juicio moral que Greene deja en el tintero. Este último es el caso, por ello merece atender de forma más amplia al autor.

En *El error de Descartes* Damasio plantea la hipótesis de los marcadores somáticos, si bien esta hipótesis comulga en líneas generales con la idea de un proceso de razonamiento y juicio moral basado tanto en las emociones como en la razón, lo cierto es que no hace una distinción de tipos de juicio tan tajante como la que hace Greene. Además otorga una complejidad de fondo al proceso de juicio moral de la que Greene carece.

Damasio plantea la insuficiencia de los procesos psicológicos puramente racionales a la hora de decidir. Un razonamiento puramente racional dirigido a la decisión debería de mantener de forma consciente en la memoria una infinidad de datos para hacer

cálculos de coste-beneficio y barajar una serie de opciones enorme sobre qué cursos de acción seguir. Sin embargo, las decisiones se toman a menudo en lapsos de tiempo muy cortos, siendo muchas de ellas descartadas de forma rápida. Esto no es posible si la decisión es procesada tan solo racionalmente. Otros factores han de entrar en juego. En este sentido Damasio reivindica el papel de las emociones en los procesos de razonamiento, en concreto cabe destacar la noción de *marcador somático*, el cual interviene de forma decisiva en la toma de decisiones, incluso en los procesos marcadamente racionales implicados en el juicio moral.

El marcador somático tiene un carácter emocional y funciona como un indicador que incentiva o inhibe determinados cursos de acción, en función de predicciones relativas a resultados considerados buenos o malos asociados a esos cursos de acción.

En resumen, los marcadores somáticos son un caso especial de sentimientos generados a partir de emociones secundarias. Estas emociones y sentimientos han sido conectados, mediante aprendizaje, a resultados futuros predecibles de determinados supuestos. Cuando un marcador somático negativo se yuxtapone a un determinado resultado futuro, la combinación funciona como un timbre de alarma. En cambio, cuando lo que se superpone es un marcador somático positivo, se convierte en una guía de incentivo. (Damasio, 2008: 205)

A simple vista puede parecer que la hipótesis del marcador somático complementa perfectamente las tesis de Greene, pero no es así. En primer lugar, Damasio considera que el marcador somático puede aparecer de forma manifiesta y de forma encubierta, en el primero de los casos se da de forma consciente en procesos de razonamiento y decisión conscientes, es decir, puede operar como una razón en procesos conscientes aun si es de corte emocional, algo que Greene no contempla, en el segundo caso funciona de forma análoga a las intuiciones tal y como Greene las describe. En segundo lugar, mientras que Damasio plantea el marcador somático como un factor emocional conectado por la experiencia a determinados resultados, Greene plantea la intuición generalmente como una predisposición biológica.

El planteamiento de Damasio tiene la carencia de no nombrar también las predisposiciones biológicas que pueden funcionar como marcador somático en forma de intuición, en el sentido que Greene plantea. Y viceversa, Greene desatiende la importancia de la experiencia y del aprendizaje, que puede generar intuiciones de forma adquirida, incluso como resultado de procesos de razonamiento y juicio anteriores, que después

pasan a interiorizarse y funcionar como intuiciones o marcadores somáticos. En este sentido la propuesta de Greene es deficitaria, porque este hecho afecta directamente a su planteamiento de la cuestión, si un marcador somático o cualquier intuición que funcione como alarma, tal y como él las ilustra, puede tener como origen la interiorización de resultados de procesos cognitivo-rationales previos, entonces esa emoción, intuición o, en términos de Damasio, ese marcador somático, harán que el juicio sea cognitivo-rationale incluso desde algunos de los factores intuitivo-emocionales que lo condicionan. Damasio considera que la mayoría de los marcadores somáticos son aprendidos como emociones secundarias. A mi juicio es precisa una mayor investigación al respecto, pues algunas predisposiciones biológicas innatas configuran emociones primarias que pueden operar como marcadores somáticos o intuiciones. En cualquier caso, la propuesta de Greene queda comprometida teniendo en cuenta las tesis de Damasio.

Así pues, los marcadores somáticos se adquieren con la experiencia, bajo el control de un sistema de preferencias interno y bajo la influencia de una serie de circunstancias externas que incluyen no solo entidades y acontecimientos con los que el organismo ha de interactuar, sino también convenciones sociales y normas éticas. (Damasio, 2008: 211)

Según plantea aquí Damasio, el papel de la emoción es activo en cualquier toma de decisión, la intervención de los marcadores somáticos se hace necesaria para que el proceso de razonamiento no sea una ardua y tediosa tarea que en última instancia no sería útil, simplifican el proceso eliminando algunos cursos de acción de la ecuación y favoreciendo otros. El resultado es una toma de decisión más liviana y efectiva, emocionalmente guiada, que libra al individuo de una tarea de razonamiento demasiado rigurosa, que es, a lo sumo, inviable.

Ahora bien, permítaseme proponer que si dicha estrategia es la única de que disponemos, la racionalidad, como se ha descrito antes, no funcionará. En el mejor de los casos, nuestra decisión tomará un tiempo excesivamente largo, mucho más de lo que sería aceptable si aquel día hemos de hacer alguna otra cosa. En el peor de los casos, puede que incluso no acabemos tomando una decisión, porque nos habremos perdido en los desvíos de nuestro cálculo. (Damasio, 2008: 203)

Como puede verse, lejos de complementar la propuesta de Greene, Damasio ofrece una bien distinta. Mientras que Greene plantea una teoría del proceso dual en la que los juicios morales se hallan vinculados bien al proceso emocional o bien al racional, siendo también categorizados de forma dual, a saber, los constitutivamente emocionales y los constitutivamente cognitivo-rationales, Damasio plantea que cualquier toma de

decisiones es constitutivamente emocional y racional, sea en mayor o menor medida. Además, considera no solo que la emoción intervenga en los procesos racionales, sino que en muchos casos es necesario que lo haga para un adecuado ejercicio de la razón y la toma de decisiones. Encontramos en este punto un apoyo a la propuesta de una razón cordial (Cortina, 2009).

Pero mientras en algunas circunstancias los impulsos biológicos y la emoción pueden dar lugar a la irracionalidad, en otras son indispensables. Los impulsos biológicos y el mecanismo automático del marcador somático que se basa en ellos son esenciales para algunos comportamientos racionales, especialmente en los ámbitos personal y social, aunque pueden ser perniciosos para la toma racional de decisiones en determinadas circunstancias al crear un sesgo avasallador contra hechos objetivos o incluso al interferir con mecanismos de soporte de la toma de decisiones, tales como la memoria funcional. (Damasio, 2008: 226)

Según Damasio, los condicionamientos emocionales como los marcadores somáticos, los impulsos biológicos y las intuiciones de corte emocional pueden ser decisivos en algunos procesos de razonamiento y toma de decisiones, mientras que en otros supone un lastre¹⁴.

Greene argumentará, como tendremos ocasión de ver en el siguiente epígrafe, que determinado tipo de juicios morales basados en la intuición son normativamente menos válidos por presentar ese lastre, sin embargo no considera la posibilidad opuesta, que la emoción resulte útil para el razonamiento en tales casos. De forma inversa considera que los juicios basados en el cálculo cognitivo-racional tienen la virtud de reducir su componente emocional, desestimando el hecho, señalado aquí por Damasio —de forma afín a la propuesta cordial de Cortina (2009)—, de que incluso el razonamiento precisa de la emoción en algunos casos.

Greene ha diseñado una serie de experimentos basados en la evaluación moral de determinados dilemas, categorizados como personales e impersonales. Aunque similares entre sí, la estructura del proceso de razonamiento y decisión de estos varía notablemente entre los dilemas categorizados diferentemente, reflejando una actividad neuronal focalizada, aunque no exclusiva, en distintas zonas del cerebro, relativas a áreas de procesado emocional y cognitivo-racional, según el caso.

¹⁴ En los párrafos subsiguientes brinda algunos ejemplos (Damasio, 2008: 225-230). Para más información sobre la racionalidad de las emociones: Damasio, 2008: 235.

Mi conjetura es que los diferentes resultados en neuroimagen obtenidos por Greene no reflejan distintos tipos de juicio, sino que reflejan los polos opuestos de un mismo proceso que puede presentar la participación tanto del sistema emocional como del cognitivo-racional, siendo posible un abanico de casos en los que pueda apreciarse una diferencia gradual en la participación de unos u otros procesos. Sin embargo, los experimentos de Greene y el tipo de dilemas que este propone en los mismos condicionan los procesos de juicio, situándolos en los polos de ese abanico. El resultado es una serie de datos científicos que parecen reflejar tipos de juicio diferentes, encuadrados en la teoría del proceso dual. En realidad, según planteamos, solo existe un tipo de juicio, con diferente grado de participación de uno u otro sistema. La variación se da, por tanto, gradualmente, no absolutamente. Aquí podemos reconocer un problema de método, el uso de dilemas es problemático, no favorece una representación fiel de los hechos, simplificando la complejidad de los procesos de juicio moral (Damasio, 2005, 2008; Mikhail, 2008; Weston, 2009; Cortina, 2012a), favoreciendo una interpretación dual de los procesos morales subyacentes, en detrimento de la propuesta de un único proceso mixto (Bluhn, 2014).

Además de los marcadores somáticos, en los procesos de razonamiento y toma de decisiones intervienen una infinidad de otros factores. Resulta curioso cómo Greene simplifica en sus textos esta complejidad de fondo en el juicio moral. Damasio ofrece algunas pistas que permiten vislumbrar esta complejidad. Entre los diferentes factores que intervienen Damasio señala el “amplio despliegue de conocimiento sobre la situación”, compuesto por “imágenes correspondientes a múltiples opciones para la acción y múltiples resultados posibles”, sumado al papel del lenguaje que configura en narrativa una serie de imágenes, todo lo cual genera una “yuxtaposición de imágenes rica y diversa, que encaja con el conocimiento previamente categorizado”. Las estructuras prefrontales que presumiblemente desempeñan estas funciones han sido acuñadas por Changeux como generador de diversidad. Este generador de diversidad maneja una cantidad de conocimiento factual y lo proyecta en la consideración de las posibilidades que pueden derivar de diversos cursos de acción. Para que sea posible este proceso de despliegue de conocimiento Damasio señala dos condiciones indispensables, han de intervenir mecanismos de *atención básica* que permitan la permanencia de la imagen y la consciencia, y también han de intervenir mecanismos de *memoria funcional básica*, que permitan mantener la distinción de imágenes durante el periodo de tiempo necesario. A

su vez la atención básica se ve impulsada por mecanismos relativos al *valor básico*, que suelen estar condicionados biológicamente. (Damasio, 2008: 230-232)

Todo ello, además, opera en una creación de orden que permite el razonamiento y la toma de decisiones. En relación al orden, Damasio llama la atención sobre la necesidad de criterios para el proceso de razonamiento y toma de decisiones. Propone lo siguiente:

Ya sea que concibamos que la razón está basada en la selección automática o que se basa en una deducción lógica mediada por un sistema simbólico, o (preferiblemente) en ambas, no podemos ignorar el problema del orden. Propongo la siguiente solución: 1) Si debe crearse orden entre las posibilidades disponibles, entonces estas deben hallarse jerarquizadas. 2) Si han de jerarquizarse, entonces se precisan criterios (valores o preferencias son términos equivalentes). 3) Los marcadores somáticos suministran criterios que expresan, en cualquier momento dado, las preferencias acumulativas que hemos recibido y adquirido. (Damasio, 2008: 233)

De nuevo la relevancia de los marcadores somáticos hace su aparición, cuando todos los elementos que intervienen en el proceso de toma de decisiones interactúan se hace necesaria la intervención de algún criterio que guíe la decisión en una dirección u otra, que ordene las posibilidades, que las jerarquice. El marcador somático, configurado bien como herencia genética, bien como predisposición biológica o bien por aprendizaje, siendo esta última la opción predilecta de Damasio, hace aquí las funciones de criterio, haciendo que cualquier toma de decisión, que será racional en mayor o menor grado, precise de un componente emocional de forma indispensable, el marcador somático. De nuevo la propuesta de Greene parece sospechosa, si Damasio está en lo cierto hemos de considerar que todos los juicios, incluso los que Greene designa como cognitivo-racionales, son constitutivamente emocionales, puesto que los criterios que guían la decisión también lo son.

La teoría computacional y la complejidad del juicio.

Además de Damasio, encontramos otros autores que plantean propuestas alternativas a la de Greene en algunos puntos. Una de estas propuestas es la ofrecida por la cognición computacional en lo que concierne a la moral. Según la teoría computacional los procesos cerebrales se producen en términos mucho más amplios y complejos que los propuestos por Greene en sus experimentos. Si bien tales experimentos aportan determinada información sobre cómo se produce el juicio moral, el valor que Greene atribuye a dicha información y la interpretación que de ella hace pueden ser puestos en

cuestión. El punto de vista de la cognición computacional es más cercano a Damasio que a Greene, aportando una perspectiva con una gran riqueza en torno a la complejidad del juicio moral.

En esta corriente encontramos John Mikhail, quien articula una crítica contra Greene en un artículo titulado "Moral Cognition and Computational Theory". Mikhail comienza el artículo planteando la posibilidad de dar una explicación más adecuada del juicio moral desde la perspectiva computacional que desde la perspectiva del proceso dual de Greene, ya que esta última plantea la cuestión de forma un tanto simplista.

Greene afirma que "la respuesta emocional... predice el juicio deontológico" (Greene, p.154), pero su propia explicación de un subconjunto de los más simples y amplios estudios de estos juicios — intuiciones sobre los problemas del tranvía— en términos de una distinción personal-impersonal no es completa ni descriptivamente adecuada (Mikhail, 2002), como Greene ahora reconoce en una reveladora nota al pie. Como sugiero más adelante, una explicación más plausible de estas intuiciones sugiere que el cerebro humano contiene una "gramática moral" computacionalmente compleja (ej., Dwyer, 1999; Harman, 2000; Mikhail, 2000; Mikhail, Sorrentino y Spelke, 1998) que es análoga en ciertos aspectos a las gramáticas mentales que operan en otros dominios, como el lenguaje, la visión, la música y el reconocimiento facial (Jackendoff, 1994). Si esto es correcto, entonces el énfasis de Greene en la emoción puede estar fuera de lugar, y al menos algunos de sus argumentos puede que necesiten ser reformulados. (Mikhail, 2007: 81-82)

La propuesta de la gramática moral de Mikhail simpatiza con la de Hauser, quien además tiene vinculación con la investigación de Mikhail, y al que atenderemos en posteriores capítulos. Según Mikhail el modelo de Haidt y Greene es demasiado simple, necesita incorporar una serie de elementos que den cuenta de cómo se produce el paso de los hechos y los escenarios, en tanto que estímulos, a una interpretación de tales estímulos y una evaluación de estos en sentido moral. Dicho de otro modo, el modelo de Greene deja demasiadas cosas en el tintero, necesita ser más complejo, mientras que un modelo computacional es capaz de dar cuenta de todo ello de forma más adecuada.

Aunque cada una de estas intuiciones se desencadena por un estímulo identificable, la forma en que la mente interpreta estos patrones de hechos hipotéticos y asigna un estado deóntico a los actos que representan no es algo revelado de manera obvia por los escenarios mismos. En cambio, se debe postular un paso intermedio: un patrón de organización de algún tipo que la mente misma impone al estímulo. Por lo tanto, un modelo perceptual simple, como el que está implícito en la influyente explicación del juicio moral de Haidt (2001), no es adecuado para explicar estas intuiciones. En cambio,

como es el caso con la percepción del lenguaje (Chomsky, 1964), un modelo perceptivo adecuado debe ser más complejo. (Mikhail, 2007: 83)

Dado que el modelo perceptual de Greene carece de la complejidad necesaria para poder explicar los procesos relativos a las intuiciones morales y al razonamiento moral, la mera asociación de cierto tipo de procesos cerebrales a cierto tipo de problemas o dilemas morales dice mucho menos sobre el juicio moral de lo que Greene presume. Si bien sus experimentos aportan algunos datos científicos interesantes para la investigación, son más anecdóticos de lo que en un principio parecen ser. (Mikhail, 2007: 87)

Críticas generales a la propuesta de Greene.

Las críticas que Damasio y Mikhail vierten sobre Greene se articulan desde modelos alternativos al del proceso dual. Resulta necesario atender a otras críticas del mismo tipo, así como a aquellas que, aceptando el modelo del proceso dual, denuncian las debilidades en la propuesta de Greene. Así haremos visible el carácter problemático de esta en lo concerniente al ámbito científico. En el siguiente epígrafe atenderemos, además, a las cuestiones filosóficas que Greene plantea en su propuesta, que precisan también ser sometidas a crítica.

Uno de los aspectos problemáticos de la propuesta de Greene deriva del uso de dilemas, ya que estos no favorecen una representación fiel de los hechos. Su propuesta simplifica la complejidad de los procesos de juicio moral, cuya estructura y elementos presentan una riqueza mayor que la planteada por Greene (Damasio, 2005, 2008; Mikhail, 2008; Weston, 2009; Cortina, 2012a). Los dilemas reproducen una situación poco realista en la que la decisión se ve condicionada por muy pocos factores y la libertad de acción está muy limitada. La información que puede obtenerse, en rigor, solo facilita datos sobre el tipo de situaciones concretas de los dilemas, extrapolar esos datos a otro tipo de situaciones no es adecuado y se ha de hacer con cautela.

Es comprensible que, en aras de poder acotar un marco abarcable por el experimento, este tiende a simplificar las condiciones del mismo, haciéndolas más controlables, pero esto tiene el efecto colateral de hacerlo menos fiel a lo que ocurre en la realidad. Uno de los principales problemas en los experimentos de Greene viene asociado a este hecho. El tipo de problemas morales a los que un individuo se enfrenta en el mundo real son de una complejidad enorme, muy pocas veces presentan una solución

unívoca, y aún si se opta por el cálculo de consecuencias este no resulta tan simple y manejable como en los *trolley dilemmas*. Los dilemas de Greene presentan respuestas unívocas según se escoja una perspectiva consecuencialista o no. Y además no reproducen de forma adecuada la estructura de los problemas morales que pueden asaltar a los individuos en el día a día. Un problema presenta muchas posibilidades de decisión y acción, es complejo, y precisa una evaluación de multitud de factores para llevar a cabo una reflexión y una toma de decisión. Sin embargo, un dilema, como su propio nombre indica, plantea una situación ofreciendo dos soluciones posibles, entre las cuales se ha de escoger. Esto supone que los experimentos que Greene utiliza como herramienta para su investigación científica del juicio moral son muy deficientes, sus resultados, por tanto, ofrecen información limitada al contexto del experimento, que ha de ser interpretada con cautela si se pretende a extender a contextos de juicio moral diferentes.

Desde una perspectiva de corte educativo, Adela Cortina expresa su preocupación en relación al hecho de que la investigación científica, y por ende sus resultados y aplicaciones, estén centrando sus esfuerzos en la resolución de dilemas morales, ya que, teniendo en cuenta la crítica que acabamos de exponer, no son dilemas el tipo de problemas que los individuos encuentran en su vida.

[...] en la vida cotidiana lo que se encontrarán niños y adultos serán *situaciones morales* y, a lo sumo, *problemas morales*, y para unas y otros es para lo que importa que estén preparados, que tengan razones y emociones a la altura de lo que creemos moralmente justo y bueno, que vayan forjándose un carácter. *Educación para resolver dilemas no es educar moralmente para la vida cotidiana.* (Cortina, 2012a: 229)

Además, la comprensión del correlato neural del juicio moral se ve adulterada por el diseño de los experimentos de Greene. Como hemos explicado con anterioridad, estos no reflejan distintos tipos de juicio, sino que, debido al tipo de dilemas manejados, reflejan los polos opuestos de un mismo proceso mixto, que puede presentar la participación tanto del sistema emocional como del cognitivo-racional, siendo posible un abanico de casos en los que pueda apreciarse una diferencia gradual en la participación de uno u otro sistema. Hemos de preguntarnos si se puede hacer una división y separación de los procesos emocionales respecto a los procesos racionales. Greene presume que tal cosa sí es posible, diferenciando dos tipos de procesos mentales —ajustes automáticos y ajustes manuales— en la elaboración de juicios morales. Clasificando las diferentes respuestas en dos tipos de juicio moral, los deontológicos y los consecuencialistas. Sin

embargo, esta división es excesivamente simplista, cabría preguntarse si pueden identificarse los juicios morales deontológicos exclusivamente con procesos emocionales, y también si acaso los juicios morales consecuencialistas no introducen aspectos emocionales.

Esta cuestión nos lleva a plantear la posibilidad de que cualquier toma de decisión o juicio moral sea constitutivamente emotivo-racional. Los factores emocionales siempre intervienen, guiando el cálculo racional de forma que solo algunos elementos son manejados de forma consciente. La decisión basada solamente en la razón y producida de forma totalmente consciente es inviable, dada la cantidad de información que habría de manejarse (Damasio, 2008: 203). Un ejemplo de ello es que los procesos racionales se ven intervenidos por factores emocionales como los marcadores somáticos¹⁵. De igual modo, puede haber intuiciones con carga racional, aquí Greene desatiende la experiencia y el aprendizaje, juicios que se producen como resultado de procesos racionales y calculados pueden interiorizarse y pasar a funcionar como intuiciones, siendo intuiciones racionalmente justificadas. Por ejemplo, muchos marcadores somáticos son aprendidos como emociones secundarias (Damasio, 2008: 211).

De forma muy cercana a esta cuestión, John Mikhail (2008) ha criticado que Greene lleva a cabo una explicación incompleta al no analizar cómo las personas consiguen interpretar los estímulos en términos de características utilitaristas o deontológicas. Esto es lo que no queda claro puesto que hay respuestas emocionales que no implican una respuesta deontológica, o viceversa, respuestas calculadas que no implican una respuesta utilitarista. Más concretamente su estudio lo que muestra es que sólo algunas de las violaciones deontológicas percibidas están asociadas con una respuesta fuertemente emocional, pero en otros casos (los contraejemplos) esto se negaría.¹⁶

Por otro lado, autores como Kamm (2009), Kahane, Wiech, Shackel, Farias, Savulescu y Tracey (2012) han presentados algunos dilemas, como el de la mentira

¹⁵ El marcador somático tiene un carácter emocional y funciona como un indicador que incentiva o inhibe determinados cursos de acción en función de predicciones relativas a resultados considerados buenos o malos asociados a esos cursos de acción (Damasio 2008).

¹⁶ “Greene no ha mostrado que la respuesta emocional predice estas intuiciones en ningún sentido significativo. Más bien, sus estudios sugieren que algunas de las violaciones deontológicas percibidas se asocian con fuertes respuestas emocionales, algo que pocos dudarían o negarían.” (Mikhail, 2007: 85).

piadosa, donde lo más intuitivo es el juicio consecuencialista y no el deontológico¹⁷, a diferencia de lo que plantea Greene. Esto ha abierto una discusión entre los autores antes mencionados, por una parte, y Greene, Paxton, Ungar y Bruni, por otra. Greene no ofrece una respuesta solvente a la hora de explicar por qué el tiempo empleado en la respuesta utilitarista es menor que el destinado al juicio deontológico en el caso de las mentiras piadosas. No se puede concluir, como hace Greene, que en base al TRC (Test de Reflexión Cognitiva) los sujetos con una mayor capacidad de cálculo matemático son más utilitaristas, y por tanto no se puede asimilar el juicio moral utilitarista con el cálculo puramente matemático. Greene confunde dos factores en el juicio moral que son independientes, a saber, el contenido del juicio moral (deontológico vs. utilitarista) con lo intuitivo o contraintuitivo de dicho juicio moral.

Además, Greene pasa por alto que el utilitarismo implica una noción del bien o mal moral. Ese bien o mal moral está vinculado, además, con alguna especie de sentimiento, emoción o intuición moral. Por ello Mikhail reclama un modelo perceptual mucho más complejo que considere las reglas de conversión, las representaciones mentales inconscientes y las reglas deónticas. Así, desde el estímulo recibido hasta la respuesta elaborada en el juicio moral habría una serie de elementos contenidos en la “gramática moral” que hay que tener en cuenta¹⁸. La complejidad, extensión y abstracción del proceso desmentiría el hecho de que las intuiciones deontológicas no dependen de un razonamiento moral genuino, complejo y sofisticado (Mikhail, 2008: 90)¹⁹. Sin embargo, a juicio de Greene, es posible un tal proceso cognitivo “impersonal” más allá de la gramática moral. De hecho, Greene parte de la premisa de que pueda haber un tipo de juicio moral que se asiente pura y llanamente en el cálculo racional (de tipo matemático) sin ningún tipo de vínculo con la dimensión afectiva de los sentimientos o las emociones. Si es puro cálculo matemático, ¿por qué llamarlo “juicio moral”? Desde

¹⁷ El propio Greene (2007: 66) ya había aludido al dilema de la mentira piadosa como excepción a la regla. Esto abrió el debate acerca de si la categorización los juicios utilitaristas como más contraintuitivos viene marcada de modo artificial por una estrecha selección de los dilemas que se plantean. Por ello, a la luz de la nueva batería de dilemas, Kahane y otros sostuvieron que dilemas como el de la mentira piadosa pueden favorecer el mayor bien (opción utilitarista) pero sin embargo son intuitivos.

¹⁸ Cabe señalar que la propuesta de Mikhail sobre una gramática moral precisa un análisis detallado, puesto que se han identificado algunas críticas sobre el autor que merece la pena tener en cuenta. Sin embargo no profundizaremos en esta cuestión, ya que se aleja del tema que estamos tratando.

¹⁹ Cuando Mikhail está apelando a que nuestros juicios morales van acompañados de una “gramática moral”, ¿se está refiriendo a que no es posible una respuesta utilitarista como interpreta Greene (2008b); o más bien, a que la respuesta utilitarista implica una gramática moral determinada, donde lo válido viene determinado por un cálculo de beneficios basado en algún valor dado (en el caso del puente peatonal el valor moral de la vida humana)?

la teoría computacional del juicio moral puede extraerse una interesante crítica a la teoría del proceso dual del juicio moral de Greene. El paso aquí fundamental es cómo la mente interpreta los estímulos identificables y les asigna un estatus deóntico. Este es el principal problema que ni el intuicionismo moral de Haidt ni la teoría del proceso dual del juicio moral de Greene consiguen explicar satisfactoriamente. De hecho, Greene tiende a omitir ese paso que es un paso moralmente decisivo, pues en él se introduce el significado moral de dicho proceso.

Atendiendo a las críticas anteriores, la propuesta de Greene queda puesta en tela de juicio desde diferentes perspectivas, presentando inconsistencia en diferentes cuestiones. Partiendo de esta inconsistente propuesta científica, Greene pretende argumentar la posibilidad de extraer conclusiones normativas. El resultado está impregnado por la huella de las críticas que hemos señalado, e introduce además errores nuevos, algunos de ellos de corte filosófico. Sobre ellos centraremos nuestra crítica en el próximo epígrafe.

3.2. Errores filosóficos y reduccionistas en la teoría del proceso dual de Greene

Las investigaciones de Joshua Greene en torno al juicio moral se han convertido en uno de los ejes de debate de la neuroética. Nuestra tesis principal en el presente epígrafe será que el autor comete una serie de errores al plantear y extraer ciertas tesis filosóficas a partir de los resultados de sus investigaciones científicas. Entre estos errores se encuentra en primer lugar un uso poco riguroso de ciertos términos filosóficos como *deontológico* y *consecuencialista*, así como de la definición de las teorías éticas asociadas a estos términos, que adultera la posibilidad de extraer conclusiones filosóficas adecuadas. Otro de los errores más destacados que comete es una transposición categorial asociada a una anfibología de los términos, ya que establece una concurrencia injustificada entre el modo en que opera la naturaleza psicológica humana y el modo en que ha de operar la filosofía en los procesos de legitimación de juicios, contenidos y normas morales. A la luz de estas críticas tendremos ocasiones de atender a otros errores cometidos por el autor, destacando el sesgo naturalista y el carácter reduccionista de su propuesta.

La pretensión normativa en la propuesta de Greene.

En 2003 vio la luz “From neural ‘is’ to moral ‘ought’: what are the moral implications of neuroscientific moral psychology?”, en este artículo Greene plantea qué tipo de conclusiones morales pueden extraerse a partir de experimentos científicos. Si bien acepta la distinción que los filósofos hacen entre el *es*, del que se ocupa la ciencia, y el *debe*, del que se ocupa la ética, eso no significa, dice, que las investigaciones neurocientíficas sobre lo que *es* no tengan implicaciones significativas para las reflexiones filosóficas sobre lo que *debe ser*.

Las implicaciones de la psicología moral para la ética, según Greene explica, no son de corte normativo, proponiendo contenidos o la validez de estos, sino revelando los mecanismos neurobiológicos que subyacen a las conductas morales humanas, “veo a la ciencia ofreciendo una mirada ‘detrás de las escenas’ en la moralidad humana” (Greene, 2003: 847), lo cual nos permitirá comprender desde una perspectiva naturalista qué es eso a lo que llamamos moral. Si lo que se está planteando es que la psicología moral se limita a dar una explicación de la naturaleza de la moralidad, entonces esta idea es aceptable. Sin embargo, este *behind the scenes*, que se presenta ambiguo en este artículo, adoptará otro tono más adelante, cristalizando en artículos posteriores en la tesis de que el deontologismo y el consecuencialismo son manifestaciones filosóficas de procesos psicológicos subyacentes, en virtud de lo cual Greene tratará de deslegitimar el carácter normativo del deontologismo.

El texto clave en relación a esta cuestión es “The Secret Joke of Kant’s Soul” (2008a), cuya propuesta se verá complementada y ampliada posteriormente en “Point-and-Shoot Morality” (2014a). La tesis que pretende defender en ambos textos es que los juicios deontológicos y los consecuencialistas corresponden a dos procesos diferentes, cada uno de los cuales opera de forma subyacente en el ejercicio del consecuencialismo y el deontologismo filosófico, respectivamente. En virtud de ello Greene pretende desacreditar la validez normativa del deontologismo, que define, erróneamente, como una racionalización *post hoc* de intuiciones emocionales, sosteniendo que “en lugar de estar basado en el *razonamiento* moral, es en gran medida un ejercicio de *racionalización* moral”. A pesar del carácter empírico de sus experimentos, pretende extraer conclusiones normativas, declarando lo siguiente: “Argumentaré que si estas afirmaciones empíricas son verdaderas, pueden tener implicaciones normativas, poniendo en duda la deontología

como una escuela de pensamiento moral normativo.” (Greene, 2007a: 36) Esta tesis se deriva de diversos errores filosóficos, que serán el foco de nuestra crítica, y a los que atenderemos a continuación.

Definición poco rigurosa de deontologismo y consecuencialismo.

Uno de los primeros errores de Greene es el modo en el que define el deontologismo y el consecuencialismo. Si bien menciona algunas de las características de estos, deja muchas otras en el tintero. Del deontologismo dice que pone énfasis en las normas morales, a menudo articuladas en términos de derechos y deberes. Del consecuencialismo dice que en la decisión moral atiende a las consecuencias de la misma para determinar qué debe hacerse y qué no. Si bien esto es cierto, es una definición tan sumaria que caricaturiza cada una de las perspectivas, dejando en la sombra todo el bagaje filosófico que constituye a cada una de ellas.

Greene cuestiona que el deontologismo y el consecuencialismo, así definidos, sean meras teorías éticas o filosofías morales (“moral philosophies”). Afirma que “los términos ‘deontología’ y ‘consecuencialismo’ refieren a *clases naturales psicológicas*”, a lo cual añade que “los enfoques consecuencialistas y deontológicos de la filosofía no son tanto invenciones filosóficas como son manifestaciones filosóficas de dos patrones psicológicos dissociables”. En función de ello pretende defender que los filósofos han malentendido en qué consisten en realidad el deontologismo y el consecuencialismo, sin comprender con qué tipo de problemas están lidiando (Greene, 2008a; Timmons, 2008)²⁰. Tras ello asegura que corresponde a la ciencia resolver el problema. (Greene, 2008a)

Cabe objetar a Greene que el deontologismo filosófico no es una racionalización de juicios formados intuitiva y emocionalmente, sino que es una propuesta que permite someter esos y otros tipos de juicios morales a crítica racional, para asentir ante su validez normativa si resisten la crítica (Korsgaard; Habermas; Cortina; Apel; Kant). El hecho de partir de una intuición emocional no es un problema si desde un nivel de abstracción superior se buscan razones para asentir ante la intuición como moralmente correcta o para desecharla (Cortina, 2012a: 107). El deontologismo versa, pues, sobre los mecanismos y condiciones que permiten otorgar validez normativa a juicios y normas,

²⁰ Una crítica más detallada es llevada a cabo por Timmons: Timmons, 2008.

independientemente del origen de estos. Además, las teorías éticas de corte deontologista presentan más variedad de la que Greene hace ver, quien atiende tan solo a Kant.

Cabe añadir que Greene no justifica bien las premisas de las que parte. En primer lugar, está suponiendo que sus experimentos son concluyentes respecto a calificar un juicio como eminentemente emocional o eminentemente racional, tesis que cuenta con fuertes críticas (Damasio, 2005, 2008; Dean, 2010; Bluhn, 2014; Mihailov, 2016; Mikhail, 2008; Weston, 2009; Cortina, 2012a). En segundo lugar, presupone que el deontologismo es principalmente emocional y que el consecuencialismo es principalmente cognitivo-racional. Esta presuposición es de corte filosófico y no está en absoluto justificada. Deontologismo y consecuencialismo no son divisibles en las categorías de emocional y cognitivo-racional. Si bien Greene reconoce que ambos participan en mayor o menor de media de ambos procesos, cree que en cada uno de los enfoques filosóficos predomina uno de los procesos psicológicos, llegando incluso a referir a estos procesos como icebergs psicológicos (“psychological icebergs”) (Greene, 2008a). Desde esta categorización caricaturiza lo que el deontologismo y el consecuencialismo son.

Transposición categorial.

Greene afirma que la ciencia importa para la ética en un modo tal que de la primera pueden extraerse conclusiones normativas para la segunda. Uno de los modos en que esto es posible es referido en “Point-and-Shoot Morality” (Greene, 2014a) como “ruta indirecta”, en este artículo argumenta que según el problema que afrontemos es más conveniente recurrir a uno o a otro de los procesos. Sostendrá que la filosofía moral ha de operar únicamente mediante el proceso consciente y controlado, en virtud de lo cual, apelando a la supuesta conexión entre las intuiciones y el deontologismo, pretende desacreditar la validez normativa de este. El argumento que Greene plantea para poner en tela de juicio la validez normativa del deontologismo tendría esta forma:

P1. Hay dos procesos a la hora de formar el juicio moral, uno constitutivamente emocional y el otro constitutivamente cognitivo-racional.

P2. El proceso emocional genera intuiciones morales de forma automática y no consciente que operan en el juicio.

P3. El proceso racional y controlado tiene un carácter consciente y se basa en el cálculo de consecuencias.

P4. Podemos vincular las filosofías morales deontológicas al proceso emocional, siendo una manifestación filosófica de este proceso psicológico subyacente.

P5. Podemos vincular las filosofías morales consecuencialistas al proceso cognitivo-racional, siendo una manifestación filosófica de este proceso psicológico subyacente.

P6. Los procesos psicológicos conscientes y controlados operan en la siguiente secuencia: primero la razón opera, después el juicio es generado como conclusión.

P7. En tanto que la filosofía moral es un proceso consciente y controlado, la validez normativa de una teoría moral descansa en el hecho de que tal teoría opere en la secuencia propia de los procesos psicológicos conscientes y controlados: razón-juicio. Una operación de racionalización *post hoc*, que presenta el orden inverso en la secuencia, juicio-razón, no tiene validez normativa, puesto que el juicio parte de factores emocionales e intuitivos.

Conclusión: El deontologismo no tiene validez normativa.

Podemos considerar P1 como verdadera solo si aceptamos como válida la teoría del proceso dual en los términos que Greene la plantea, considerando como válidos los resultados de esta. No obstante aceptar la validez de esta teoría resulta problemático, como hemos tenido ocasión de ver (Damasio, 2005, 2008; Dean, 2010; Bluhn, 2014; Mikhail; Weston; Cortina, 2012a; Mihailov, 2016). Suponiendo que aceptemos su propuesta sobre el proceso dual, por las mismas razones, podemos considerar verdaderas P2, P3 y P6. En cuanto a P4 y P5 son premisas que no son verdaderas, las tesis que defienden vienen dadas por la descripción poco rigurosa y caricaturizada que Greene hace del deontologismo y del consecuencialismo, así como por un uso anfibológico y confuso de los términos, al que atenderemos más adelante de forma detallada.

En cuanto a P7, podemos identificar en ella el error lógico sobre el que descansa el argumento de Greene. Cabe preguntarse por qué este considera que un razonamiento *post hoc* de una convicción o juicio moral no es un buen método de legitimación

normativa. Aun si una convicción moral parte de una intuición basada en emociones, un proceso de razonamiento crítico puede dirimir si hay o no razones que apoyen tal convicción, y en función de ello se puede asentir ante la validez moral de tal convicción (Cortina; Habermas; Apel; Korsgaard; Kant).

Nuestra tesis es que Greene comete un error lógico de transposición categorial, en el cual se basa P7. Parte de un hecho natural, cómo funciona nuestra psicología, y trata de concluir cómo debe de operar la filosofía moral. Ese paso del *es* al *debe*, en rigor, no comete una falacia naturalista. Lo que sucede es que el esquema secuencial de elementos que de hecho opera en la psicología moral, que se ubica en la esfera natural, es tomado como criterio de validez normativa para teorías éticas que operan en la esfera de la filosofía moral. Es esta extrapolación de una esfera a otra la que conduce a la transposición categorial.

	Cómo opera la naturaleza	Cómo opera la ética
Tipo de proceso consciente y controlado	Psicológico	Filosófico
Orden en la secuencia de elementos en procesos conscientes y controlados	Razón → Juicio	Dos opciones: a) Razón→Juicio b) Juicio (intuitivo) → Razón (función crítica) *La crítica conduce a un asentimiento, un rechazo o una corrección del juicio, según el caso.
Error	Transposición categorial: Greene transfiere el esquema de la naturaleza al de la filosofía.	

Tabla 4. Transposición categorial en la propuesta de Greene.

En el orden natural el proceso psicológico consciente y controlado produce juicios que son dados en la secuencia razón-juicio. Greene cree que la filosofía moral, por ejercerse consciente y controladamente, ha de operar bajo este mismo esquema, para así evitar la intromisión de elementos intuitivos y emocionales en el juicio. Sin embargo, en filosofía la secuencia opuesta (juicio-razonamiento) también es válida como mecanismo

de legitimación normativa, si un juicio intuitivamente formado resiste la crítica racional queda legitimado moralmente. Extrapolar el orden psicológico de la secuencia en los juicios producidos consciente y controladamente, que es el propio de la esfera natural, a la esfera filosófica, hace a Greene incurrir en un error argumentativo por transposición categorial. Este error le lleva a plantear que el deontologismo filosófico no es normativamente significativo y que la única alternativa que cumple esas condiciones es el consecuencialismo, que sí opera en la secuencia razón-juicio, según plantea Greene²¹. En *Moral Tribes*, retomando esta convicción, planteará que el único modo de razonamiento moral adecuado es el llevado a cabo por el consecuencialismo en su versión del “pragmatismo profundo” (Greene, 2013).

Anfibología de los términos.

El hecho de establecer un nexo entre ambas esferas hace que Greene se apropie de los términos filosóficos *deontológico* y *consecuencialista*. Al acuñar así los dos tipos de juicio que él diferencia, que se situarían en el ámbito natural de la psicología, y no en el filosófico, no está haciendo un uso riguroso de los términos filosóficos, sino que está creando dos términos homónimos, pero no sinónimos. Este hecho viene asociado a una serie de confusiones, entre ellas la transposición categorial, asociada a una anfibología en los términos, en la cual profundizaremos a continuación, así como a una falacia de ambigüedad por equívoco.

La propuesta de Greene, en todo caso, tendría sentido como un informe científico acerca de la naturaleza de juicios que deberían llamarse no consecuencialistas, en vez de deontológicos. En “The Secret Joke of Kant’s Soul” parece reconocer el aspecto problemático de su terminología, afirmando lo siguiente: “me he referido a estas intuiciones y los juicios que sostienen como ‘deontológicos’, pero quizás sería más apropiado llamarlos no-consecuencialistas” (Greene, 2008a). Más tarde, en “Point-and-Shoot Morality”, artículo en el que se muestra más cauteloso a la hora de usar una terminología poco rigurosa, toma la precaución de aclarar qué entiende él al usar estos términos, reconociendo lo siguiente: “mi uso de ‘deontológico’ y ‘consecuencialista’ no es enteramente coherente con el uso filosófico estándar, lo que exige mi uso del

²¹ En realidad el consecuencialismo, como ya hemos indicado anteriormente, sí incluye factores emocionales e intuitivos, especialmente a la hora de valorar qué consecuencias son tomadas como buenas o malas.

calificador técnico ‘característicamente’.” (Greene, 2014a: 7-8) A pesar de reconocer que el uso que hace de tales términos no es enteramente consistente con el uso filosófico estándar insiste en seguir usando esta terminología anfibológica y conducente a confusión²², en virtud de lo cual se produce una falacia de ambigüedad por equívoco conducente a un error argumentativo por transposición categorial.

Si la cuestión se reduce a un problema terminológico podría aceptarse, sea o no la opción más acertada, que Greene adopte unos términos filosóficos para un uso que le es útil en psicología moral, pero esto supone que los términos serán homónimos, es decir, significarán cosas diferentes, aunque tengan la misma forma. En tal caso, este carácter no sinonímico deshabilita el uso que Greene hace de los términos, al atender a datos científicos, para extraer conclusiones filosóficas y normativas en relación al deontologismo y el consecuencialismo.

Greene afirma que " definiendo ‘consecuencialismo’ y ‘deontología’ en términos de sus juicios característicos, le damos a nuestra hipótesis empírica una oportunidad" (Greene, 2008a), se refiere a una oportunidad de ser normativamente relevante. Ahora bien, tal y como él indica, esta oportunidad se produce en virtud de definir consecuencialismo y deontologismo en los términos de sus juicios característicos. Aquí precisamente se está produciendo el salto entre las esferas psicológica natural y la esfera filosófica, asociado al error por transposición categorial. No puede definirse el consecuencialismo en virtud de sus juicios característicos porque no existen estos juicios como tales, los juicios llamados deontológicos y consecuencialistas hacen un uso no sinonímico de estos términos respecto a sus homónimos en el campo de la filosofía. Por tanto, las conclusiones filosóficas que Greene plantea son erróneas. Además, existe una contradicción en el planteamiento terminológico, por un lado Greene afirma que los términos *juicio deontológico* y *juicio consecuencialista* son acuñados para un uso restringido al ámbito de la psicología moral, y sin embargo, en virtud de esos términos, pretende extraer conclusiones filosóficas en los ámbitos propios del deontologismo y el consecuencialismo filosóficos.

²² Otros términos también son usados anfibológicamente, en relación a “los problemas personales” no llega a aclarar bien en qué consisten, confundiendo diferentes esferas, a saber, la de la destreza técnica con la de otros problemas del comportamiento, como la conducta moral, y el plano del desarrollo evolutivo con el de la corrección o incorrección moral.

The naturalistic joke of Greene's Soul.

Los errores filosóficos que hemos denunciado en el apartado anterior denotan un claro sesgo naturalista en Greene. La definición deficitaria del deontologismo y el consecuencialismo, el carácter anfibológico de su terminología, la falacia de ambigüedad por equívoco y el error de transposición categorial muestran una pretendida reducción del deontologismo y el consecuencialismo al ámbito de lo natural, concretamente a dos procesos psicológicos naturales diferenciados y vinculados al juicio moral. Es preciso someter a crítica este carácter de reduccionismo naturalista que su propuesta presenta²³.

La naturalización del deontologismo y el consecuencialismo en Greene.

Al plantear una correlación entre los juicios basados en emociones y el deontologismo, por un lado, y los juicios basados en un cálculo racional y el consecuencialismo, por otro, Greene está naturalizando estas teorías éticas. Concretamente pretende naturalizar el deontologismo al incidir en que los juicios morales deontológicos se basan únicamente en respuestas emocionales e intuitivas, cuya actividad se podría localizar principalmente en la amígdala, la corteza prefrontal media, la corteza prefrontal ventromedial, el lóbulo parietal inferior y el surco temporal posterior superior (Greene, 2001: 2106-2107; Greene et al., 2004: 391-392; Greene, 2009). Análogamente pretende naturalizar el consecuencialismo vinculándolo únicamente al cálculo controlado y razonado de consecuencias, que estaría localizado en la corteza dorsolateral prefrontal. Esta pretensión reduccionista se hace aún más evidente cuando el autor defiende la tesis de que el deontologismo no es más que una manifestación filosófica de los estados y procesos psicológicos vinculados a la emoción y la intuición, tratando de negar, en virtud de ello, su relevancia normativa. Asimismo, considera que el consecuencialismo es la manifestación filosófica de patrones psicológicos subyacentes de razonamiento controlado y consciente. Pretender naturalizar el deontologismo conduce a una perversión de este, ya que —como acaba admitiendo Greene— esta postura “define la deontología como valores que no son distintivamente deontológicos” (Greene, 2008a: 73).

Greene llega a reconocer que sería más apropiado hablar de “no consecuencialista” que de “deontologista” (Greene, 2008a: 75), mostrando así que su

²³ Berker lleva a cabo una crítica de estas características en su artículo “The Normative insignificance of Neuroscience” (Berker, 2009).

definición de deontologismo no es rigurosa con el significado del término. Más bien, su terminología correspondería a una versión deficitaria de la deontología en el marco naturalista, de forma que resulte útil para satisfacer las necesidades de su investigación. Mediante este inadecuado uso de los términos, Greene nos conduce a un enfoque naturalista que plantea las emociones como el fundamento del deber, confundiendo los motivos de la acción, que pueden ser empíricos y naturales, con el fundamento normativo, que no puede reducirse a la esfera de lo natural (Korsgaard, 2000; Apel, 1985; Habermas, 2006; Cortina, 2012a; Evers, 2010; Gould, 1999; García-Marzá, 2016).

Neurociencia normativa y el sesgo naturalista de Greene.

Para no incurrir en la falacia naturalista que confunde el “es” neuronal y el “debe” moral es importante distinguir entre “bases neuronales” y “fundamento moral”. Mientras que las bases neuronales apelan al sustrato fisiológico del cerebro que capacita al ser humano para el juicio moral, el fundamento moral refiere a los procedimientos y razón que permiten justificar normas, valores y virtudes morales (Cortina, 2012a: 96). Considerar la existencia de un fundamento moral en el cerebro permitiría deducir a partir del estudio de este una serie de contenidos y normas morales, como algunos autores plantean (Casebeer, 2003; Churchland, 2013; Mora, 2007; Gazzaniga, 2006; Greene, 2013), esta pretensión incurre en reduccionismo naturalista (Buchanan, 2009; Apel, 1985; Habermas, 2006; Aranguren, 1994; Gould, 1999; Hume 1988, 1953; Moore, 1983; Salles, 2013, 2016, 2018; Evers, 2010; Cartwright, 2010; Teehan y diCarlo, 2004; García-Marzá, 2016; Conill, 2012).

Efectivamente la neurociencia nos permite conocer qué mecanismos de nuestro cerebro son accionados cuando reaccionamos de una manera u otra, cuando actuamos de cierto modo y cuando hacemos un juicio moral. Sin embargo, este conocimiento no permite extraer conclusiones sobre *lo que debe valer o no moralmente*. De hecho, más bien lo que prueba es que hay una importante disonancia y un salto cualitativo entre las tendencias naturales a realizar un juicio moral, por una parte, y lo que debe valer moralmente, por otra. Greene, sin embargo, sostiene que “hechos científicos interesantes contribuyen a alcanzar conclusiones normativas interesantes”, abogando por una “psicología moral con relevancia para la ética, esto es, significativa normativamente” (Greene, 2014a: 711-713).

Las asunciones normativas no empíricas no son tan poco interesantes como Greene quiere hacer creer (Greene, 2014: 711). Dichas asunciones normativas merecen mucha más atención y es en ellas donde se dirime la validez moral de determinados juicios. No es la neurociencia la que introduce qué es moralmente relevante, de hecho, por sí misma, esta es incapaz de determinar qué debe valer y qué no debe valer moralmente. Más bien son las asunciones normativas que emergen de otras instancias, concretamente de la filosofía moral, las que permiten ofrecer un criterio acerca de la validez normativa y moral. Son estas cuestiones morales normativamente sustantivas las que Greene tiende a omitir. Si bien, en rigor, Greene no está incurriendo en falacia naturalista, sí comete errores de magnitud equiparable, conducentes a un reduccionismo naturalista.

PARTE II: NEUROÉTICA, ENTRE LA CIENCIA Y LA FILOSOFÍA MORAL

En todo sistema moral de que haya tenido noticia, hasta ahora, he podido siempre observar que el autor sigue durante cierto tiempo el modo de hablar ordinario, estableciendo la existencia de Dios o realizando observaciones sobre quehaceres humanos, y, de pronto, me encuentro con la sorpresa de que, en vez de las cópulas habituales de las proposiciones: es y no es, no veo ninguna proposición que no esté conectada con un debe o un no debe.

David Hume, *Tratado de la naturaleza humana*

En los capítulos precedentes hemos hecho una aproximación al modo en que el estudio de la dimensión natural de la moral puede llevarse a cabo desde las ciencias empíricas. Algunos problemas de método han sido ya expuestos, especialmente en torno al problema del juicio moral. Todo ello permite dibujar el contexto en el cual la neuroética encuentra algunos de sus principales retos metodológicos. Este contexto se ve marcado por una interacción entre ciencia y ética acompañada por ciertas tensiones difíciles de resolver. Es necesario profundizar en la comprensión del estudio científico de la moralidad y de su relación con el quehacer de la ética. En las próximas páginas nos disponemos a analizar algunos de estos problemáticos aspectos en torno al método en neuroética. Trataremos de ofrecer criterios para ofrecer respuestas adecuadas, que serán clave para nuestro objetivo en la investigación, a saber, confeccionar un método adecuado para una neuroética crítica en clave cordial.

En primer lugar, atenderemos a la distinción entre estructura y contenido en lo moral, así como los errores metodológicos derivados de una mala comprensión de dicha distinción. Tras ello abordaremos la cuestión del libre albedrío y del estudio neurocientífico de este, intentando ofrecer algunas indicaciones metodológicas al respecto. Tras ello, atenderemos a dos propuestas metodológicas para el ejercicio de la neuroética, la neurofilosofía práctica de Churchland, marcadamente naturalista, y el materialismo ilustrado de Evers, que permite complementar ciencia y ética de forma adecuada. Contraponer ambas propuestas permitirá extraer algunas conclusiones interesantes a nivel metodológico. Finalmente, a la luz de las diferentes cuestiones tratadas hasta el momento, haremos un análisis de los principales errores lógicos y argumentativos cometidos por el

reduccionismo naturalista, siendo este uno de los principales problemas metodológicos presentes en la neuroética.

Capítulo 4. Moral como estructura y moral como contenido: consideraciones metodológicas

La neuroética es una rama joven del conocimiento, en la que operan conjuntamente la ciencia y la filosofía. Por ello es preciso estar alerta ante los riesgos metodológicos que esta interdisciplinariedad puede suponer. En este capítulo atenderemos a la distinción entre moral como estructura y moral como contenido, lo cual permitirá realizar una serie de recomendaciones metodológicas mediante un estudio de las propuestas más relevantes para la cuestión. En primer lugar, atenderemos a los conceptos de estructura y contenidos, así como a los errores asociados a la confusión o asimilación de estos conceptos. En segundo lugar, atenderemos a la cuestión si existe el libre albedrío, entendido como hecho natural y como una de las estructuras de la moralidad, preguntándonos si las ciencias empíricas, y en concreto las neurociencias, pueden abordar su estudio y, en el caso de que así sea, cómo debe hacerlo.

4.1. Estructuras, contenidos y fundamentación de lo moral: el error reduccionista de asimilar bases y fundamentos

A lo largo del epígrafe analizaremos uno de los errores metodológicos que algunas propuestas de neuroética han cometido, conduciendo a un reduccionismo naturalista. Este error consiste en la confusión y asimilación de las estructuras neuronales de la conducta moral con los contenidos y fundamentos normativos de esta. El resultado de ello es la pretensión reduccionista de fundamentar la ética en el cerebro, correspondiendo a la neurociencia la tarea de descubrir qué códigos normativos han de regir nuestra conducta. Para evitar este error se hace necesaria la distinción entre base cerebral y fundamentación ética de las conductas morales, cuyo estudio corresponde a dos disciplinas con diferentes metodologías, a saber, la ciencia en el primer caso, y la filosofía moral en el segundo.

En la dimensión moral del ser humano interactúan dos dominios que se complementan y que la constituyen, estos son la biología y la cultura. Existen unas estructuras psicológicas, neuronales y biológicas que permiten, en tanto que condición de posibilidad, la adquisición de diferentes normas, criterios, caracteres y conceptos morales, que serán contingentes e históricos, variando entre individuos, culturas y sociedades. La moral queda así estratificada en estructuras y contenidos. El primer estrato corresponde a

una serie de elementos naturales cuyo estudio puede ser abordado por las ciencias empíricas desde su paradigma. Sin embargo, en lo que refiere a la fundamentación de los contenidos morales la ciencia es infecunda, siendo necesario recurrir a la labor de la filosofía moral (Cortina, 2012a; Salles, 2013, 2016, 2018; Paris Albert y Comins Mingol, 2013; Habermas, 2006; Apel, 1985; García-Marzá, 2016). El carácter contingente de los contenidos no significa que debamos aceptar un relativismo moral en relación a los mismos, precisamente trataremos de especificar los criterios mediante los cuales es posible una fundamentación con pretensión de universalidad. Ciertos sectores de la comunidad científica sostienen que es posible resolver los problemas de fundamentación normativa desde el paradigma naturalista (Spencer 1964, 1883; Richards, 1986; Arnhart 1998; Singer, 1981; Casebeer 2003, 34; Flanagan; Cartwright 2000, 328; McShea y McShea 1990, 1999; Churchland 2013; Curry 2006, 235, 237-238; Mora, 2007; Gazzaniga, 2006; Greene, 2013). Esta pretensión descansa en un error, cometiendo un reduccionismo que no resiste la crítica (Buchanan, 2009; Apel, 1985; Habermas, 2006; Aranguren, 1994; Gould, 1999; Hume 1988, 1953; Moore, 1983; Salles, 2013, 2016, 2018; Jácomo, 2011, 2013a, 2013b, 2014; Evers, 2010; Cartwright, 2010; Teehan y diCarlo, 2004; Cortina, 2012a; García-Marzá, 2016; Conill, 2012). Las causas de este error se deben a una confusión entre estructuras y contenidos (Cortina, 2012a), consistente en la idea de que las predisposiciones condicionadas por las estructuras psicobiológicas de la moral implican una serie de conclusiones acerca de la validez normativa de ciertos contenidos. Nos proponemos hacer un análisis del problema atendiendo a tres cuestiones. En primer lugar, analizaremos los conceptos de moral como estructura y moral como contenido, así como la relación entre ambos. En segundo lugar, profundizaremos en esa relación desde una doble perspectiva, por un lado abordándola desde la filogénesis de nuestra especie y por otro desde la ontogénesis. En tercer lugar, analizaremos algunos de los casos en que se comete el reduccionismo naturalista mencionado, tratando de ofrecer algunas claves acerca de los motivos de este error y cómo evitarlo.

Moral como estructura y moral como contenido.

La primera de las tareas que nos proponemos realizar es hacer una aclaración de los conceptos de *moral como estructura* y *moral como contenido*. Para ello atenderemos a Aranguren. En el séptimo capítulo de su *Ética*, titulado “La realidad constitutivamente

moral del hombre: moral como estructura” encontramos las claves para distinguir estos conceptos. Además, haremos una revisión de algunos aspectos de su propuesta que, considero, pueden ser actualizados a la luz de los avances que la psicología cognitiva y las neurociencias han aportado con posteridad a la publicación de su texto.

Estructuras conductuales e inespecificidad de la respuesta.

Para comprender qué hace constitutivamente moral al ser humano resulta muy útil analizar qué elementos de su conducta pertenecen a la especie de forma exclusiva, ya que solo el ser humano es un agente moral en el pleno sentido del término. Por ello Aranguren comienza el capítulo distinguiendo el comportamiento humano del comportamiento del resto de animales. Respecto a su propuesta en esta cuestión haremos algunas correcciones que considero pertinentes. Veamos cómo caracteriza la conducta animal y qué tenemos que objetar al respecto.

En el animal, la situación estimulante de un lado y sus propias capacidades biológicas del otro, determinan unívocamente una respuesta o una serie de respuestas que establecen y restablecen un equilibrio dinámico. Los estímulos suscitan respuestas en principio perfectamente adecuadas siempre a aquellos. Hay así un «ajustamiento» perfecto, una determinación *ad unum* entre el animal y su medio. Al carácter formal de este ajustamiento le llama Zubiri «justeza». (Aranguren, 1994: 47)

El esquema estímulo-respuesta según el cual Aranguren y Zubiri caracterizan la conducta animal presenta un aspecto que es criticable. La psicología cognitiva, la etología y las neurociencias han revelado que la conducta de muchos animales, si bien no necesariamente de todos, se puede ajustar a una estructura con tres elementos, que conformarían la secuencia estímulo-cognición-respuesta (Damasio; Tomasello; Waal), siendo este un esquema más adecuado para comprender tanto la conducta humana como la de algunas otras especies animales. Lo que diferenciará al humano del resto de animales, según proponemos, no será la ausencia de cognición en otras especies al ofrecer una respuesta al medio, sino el tipo de cognición que medie entre el estímulo y la respuesta.

En cuanto a la conducta humana, Aranguren la caracteriza del siguiente modo:

El hombre comparte parcialmente esta condición. Pero el organismo humano, a fuerza de complicación y formalización, no puede ya dar, en todos los casos, por sí mismo, respuesta

adecuada o ajustada, y queda así en suspenso ante los estímulos, «libre-de» ellos (primera dimensión de la libertad). Las estructuras somáticas *exigen* así, para la viabilidad de este hiperformalizado ser vivo, la aparición de la inteligencia. ¿Qué significa aquí, primariamente, la palabra «inteligencia»? Pura y simplemente que el hombre, para subsistir biológicamente, necesita «hacerse cargo» de la situación, habérselas (concepto de «habitud») con las cosas —y consigo mismo— como «realidad» y no meramente como estímulos. (Aranguren, 1994: 47)

El ser humano, caracterizado así, ante un estímulo no tiene una respuesta predeterminada por su biología, en ese sentido es *libre-de* los estímulos. Para configurar una respuesta ha de hacer uso de la inteligencia. No obstante, tal y como hemos indicado, este aspecto no es exclusivo de nuestra especie, otros animales pueden presentar también la cognición como mediación entre estímulo y respuesta, por ello los miembros de esas especies estarán en cierto sentido libres de los estímulos, no estando su conducta determinada unívocamente por esos estímulos que ofrece el medio. Así pues, compartimos con otras especies animales la existencia de unas tendencias —o «ferencias» en la terminología de Aranguren— inconclusas, en forma de predisposiciones biológicas que orientan la acción probabilísticamente, y no deterministamente. Será la cognición la que complete el proceso de respuesta al medio. La elección entre una posibilidad u otra pone en ejercicio la otra dimensión de la libertad, la *libertad-para* preferir una acción y no otra (Aranguren, 1994: 48).

El hecho de considerar la realidad antes de actuar, por tanto, no es algo exclusivo de nuestra especie. Sin embargo, nuestra especie sí tiene formas exclusivas de considerar la realidad. En relación a los motivos que conducen a una decisión y una acción concretas, entre las diferentes posibles, Aranguren nos habla de la necesidad de justificación. Esta puede producirse en relación al nivel estructural, justificación como ajustamiento, y en relación a un nivel metacognitivo, justificación como justicia.

Justificación como ajustamiento y justificación como justicia.

Uno de los sentidos posibles de la justificación es el de *justificación como ajustamiento*, correspondiente a la dimensión estructural de la conducta, este tipo de justificación es compartida por humanos y otros animales.

Todo acto para ser verdaderamente humano tiene que ser «justo», es decir, ajustado a la realidad, coherente con ella, «respondiente a ella». Pero entiéndase bien, se trata de una primera dimensión

de la «justificación», *justificación como ajustamiento*. A esta primera dimensión la llamaremos, con Zubiri, *moral como estructura*. (Aranguren, 1994: 49)

Cuando el esquema conductual no sigue la secuencia estímulo-respuesta sino estímulo-cognición-respuesta la conducta puede ser justificada en sentido de ajustamiento, es decir, es una justificación concerniente al proceso cognitivo de responder al medio. Podemos hablar de justificación como ajustamiento cognitivo tanto en animales como en humanos, sin que necesariamente signifique que los animales también tengan moral como estructura. Para que las estructuras de ajustamiento conductual sean, además, estructuras de la moralidad es preciso que exista el otro nivel de justificación, y, por lo que sabemos, este se da solo en nuestra especie.

Este segundo nivel de la justificación se produce cuando existen unas facultades intelectuales que permiten la metacognición, haciendo posible preguntarnos por los motivos de la elección. La metacognición permite tomar el proceso de cognición como objeto de justificación. Encontramos aquí al individuo no enfrentado al medio para ofrecer una respuesta (ajustamiento) sino enfrentado a los procesos cognitivos de elección, con el objetivo de dar razones que justifiquen su corrección en un sentido moral (justicia).

La justificación puede tener un segundo sentido, al que apenas se ha hecho más que aludir al final de nuestra exposición, *justificación como justicia*. Consiste en que el acto se ajuste, no ya a la situación, no ya a la realidad, sino a la norma ética (fin último, ley natural, conciencia moral). «Justo» ahora ya no significa simplemente «ajustado», sino que significa también «honesto» [...] [Aranguren usa aquí el término *honesto* como sinónimo de *justicia*, como *facere bonum*]. A esta segunda dimensión [...] la llamaremos, con Zubiri, *moral como contenido*. (Aranguren, 1994: 49)

Esta segunda acepción de justificación ya no refiere solamente al proceso de ajustamiento a la realidad, sino a las razones que pueden darse de ese proceso. La moral como contenido, pues, corresponde en un sentido ético y reflexivo a la justificación de una conducta como justa, como moralmente correcta, en un sentido moral más amplio corresponde además a todos aquellos contenidos que no vienen dados por las estructuras biológicas, pero que completan la inespecificidad de estas, haciendo que la conducta tome una forma y no otra entre las distintas posibilidades.

Los diferentes contenidos morales precisan la existencia de las estructuras morales como soporte sobre el que asentarse, constituyendo estas sus condiciones de posibilidad.

De las dos dimensiones de la moral, la moral como contenido se monta necesariamente sobre la moral como estructura y no puede darse sin ella. Precisamente porque al hombre no le es dado por naturaleza el ajustamiento a la realidad, sino que tiene que hacerlo por sí mismo, cobra sentido demandarle que lo haga, no arbitraria o subjetivamente, sino conforme a determinadas normas, conforme a determinados sistemas de preferencias. (Aranguren, 1994: 50)

Puesto que el ser humano ha de completar el ajustamiento por sí mismo tiene sentido demandar que lo haga en virtud de unos criterios, y no arbitrariamente. Esta es precisamente la operación de la ética, que consiste en argumentar el porqué de la elección de una posibilidad concreta a la hora de actuar, dando razón de ello. Elegir sin justificar —en el segundo sentido de justificación— es posible, tal es el caso en otras especies, y también eventualmente en la nuestra. La conducta humana de elección sin justificación puede ser considerada como una conducta moral, a pesar de no justificarse en el sentido ético, puesto que el ser humano es estructuralmente moral en el primero de los sentidos de justificación, aun así no sería ética en tanto que no se ha producido una justificación metacognitiva de la elección, es decir, no se ha dado justificación en el segundo de los sentidos. Para que esto no suceda Aranguren exige la necesidad de justificar la elección en el segundo sentido de justificación, que es el genuinamente ético, evitando así la arbitrariedad y/o la subjetividad en la elección. Este segundo sentido de justificación exige dar razones que permitan justificar la elección en sentido moral, produciéndose esto desde el ámbito de la reflexión filosófica concerniente a la fundamentación ética.

Psicobiología de la estratificación de la moral.

La segunda tarea que nos proponemos llevar a cabo es facilitar una comprensión más profunda de los conceptos de moral como estructura y de moral como contenido, para ello atenderemos al desarrollo filogenético y ontogenético de las estructuras psicobiológicas que condicionan la respuesta al medio, sin determinarla, dejando espacio para que los caracteres culturales que componen los diferentes contenidos morales completen el proceso de respuesta en la decisión y en la acción.

Filogénesis de la estratificación moral.

El desarrollo cognitivo en la *filogénesis* de las diferentes especies se ve marcado por dos conjuntos de elementos interrelacionados entre sí: genética y cultura. Estos dos acervos, genético y cultural, son variables y están en continua evolución, marcada por la

interacción entre ambos (Ayala y Cela Conde, 2001: 530-532). La dotación genética garantiza una serie de adaptaciones al medio, pero presenta un problema en tanto que determine las conductas de forma cuasi mecánica o unívoca. La *determinación genética* puede ser provechosa en casos en los que el medio es estable y los problemas de adecuación pueden solventarse fijando programas de conducta en sentido determinista a nivel genético y biológico. Pero en tanto que el medio presenta elementos variables, la resolución de problemas y la adecuación al medio se ve enormemente limitada desde una estrategia evolutiva que determine la conducta del individuo. Es en este sentido en el que toma importancia el papel de la cultura, entendida como acervo de conductas adquiridas por ensayo y error y heredadas por aprendizaje, bien mediante imitación o mediante enseñanza focalizada entre miembros del grupo. La existencia de un conjunto de conductas adquiridas por aprendizaje y el desarrollo de una serie de capacidades cognitivas que permitan actualizar ese acervo de conductas supone una configuración conductual que aporta la versatilidad necesaria para adaptarse a un medio estable en términos generales pero variable y cambiante en algunos aspectos locales.

Así pues, la determinación genética es desplazada en algunas especies en favor del *condicionamiento genético*, este consiste en la existencia de ciertas predisposiciones que favorecen probabilísticamente cierto tipo de conductas, sin determinarlas. Ante los problemas de un determinismo o una libertad absolutos el condicionamiento genético ofrece una solución muy rentable, los genes favorecen un marco dentro del cual tiene cabida un amplio espectro de conductas sujetas a contingencia y adquiridas culturalmente, que permiten una adaptación rápida y ajustada a elementos variables del medio (Ayala y Cela Conde, 2001: 530-532). Dada la versatilidad de esas variables del medio no existe la posibilidad de una adaptación genética que programe unívocamente ciertas conductas ante ciertos estímulos, pues estos pueden variar. Una adaptación determinista precisa evolucionar durante varias generaciones, lo cual a menudo supone un lapso de tiempo muy superior al periodo de duración de ciertos cambios del medio, sujetos a variables, así como al de los problemas que estos implican para los miembros de las diferentes especies. La previsión genética ante esta situación no se ha configurado en forma de programación unívoca y determinista sino en forma de condicionamiento orientativo. Dicho de otro modo, genéticamente se han desarrollado unas estructuras capaces de generar y albergar diferentes contenidos que se adecúen a los diferentes problemas y variaciones de un medio que a gran escala parece estático, pero que en el seno de ese estatismo muestra

aspectos dinámicos y cambiantes. En el sentido que Aranguren ilustra la cuestión, hablaríamos aquí de estar *libre-de*, en la medida en que los estímulos no determinan la respuesta, y *libre-para*, en tanto que existen diferentes respuestas posibles ante un mismo estímulo, todo ello en un marco que sí viene dado biológicamente, pero que tan solo es orientativo, en su terminología hablaríamos aquí de las *ferencias* (Aranguren, 1994).

Estas adaptaciones que predisponen probabilísticamente tienen alcance en la esfera de las conductas sociales y morales. Hablar de moral como estructura y de moral como contenido hace referencia a este tipo de condicionamiento genético, en el que una estructura biológica orienta probabilísticamente la conducta dejando espacio para contenidos de carácter cultural y contingente.

No puede hablarse en ese caso de un «código ético universal», sino de una «tendencia universal a la aceptación de códigos éticos». Si aceptamos esta última perspectiva, la relación existente entre control genético y conducta moral debe entenderse, por lo que hace al altruismo de la especie humana al menos, en un sentido débil. Nos situamos así muy cerca de las posturas neodarwinistas, en las que ya se mantenía que los genes limitan su control al mero establecimiento de disposiciones hacia la acción moral. (Ayala y Cela Conde, 2001: 533)

La interrelación entre genes y cultura no se agota en esta cuestión. Las conductas culturales adquiridas se prolongan en el tiempo y en las generaciones, condicionando la futura evolución biológica. Por ejemplo, una vez adquirida la conducta de abrigarse, la presencia de pelo o grasa dejó de operar como elemento de selección, teniendo éxito evolutivo individuos biológicamente menos preparados para el frío, pero que contaban con una solución a este problema adquirida culturalmente. Esto significa que la resolución cultural de determinados problemas del medio, al ser más versátil y rápida que la adaptación genética, en numerosos casos la sustituye, inhibiendo posibles factores de selección. Se produce así una inhibición de adaptaciones genéticas a problemas específicos en la medida en que prospere una cognición conductual que predisponga probabilísticamente. Es decir, se favorece el condicionamiento genético frente al determinismo genético.

Cada individuo supone una actualización única e irrepetible de estos dos conjuntos, genético y cultural, configurada por su dotación genética y por el proceso ontogenético de endoculturación. Si hemos de buscar una naturaleza moral humana, la mejor definición parece ser aquella que reconoce a un mismo tiempo su condicionamiento

estructural y su indeterminación en cuanto a contenidos. En la medida en que el ser humano tenga la capacidad de analizar y evaluar los diferentes contenidos morales, y en la medida en que sea capaz de elegir libremente unos u otros contenidos, esa indeterminación moral puede transformarse en autodeterminación, siendo esta necesaria para una conducta verdaderamente libre. Entra aquí en juego el problema metacognitivo de la justificación de la acción, que constituye al ser humano como agente moral. Sobre el problema de si existe o no el libre albedrío en el ser humano, es una cuestión que no ha podido ser resuelta a día de hoy, a pesar de que en las últimas décadas la neurociencia ha puesto el foco de algunas de sus investigaciones en este problema, la cuestión que cabe preguntarse es si la tarea de resolver esta cuestión está al alcance de las ciencias empíricas y su método (Cortina, 2012a; Evers, 2010; Libet 1999; Roskies, 2010; Ortiz de Landázuri, 2016, Pallarés-Domínguez, 2013, 2016), atenderemos a esta cuestión de forma detallada en el siguiente epígrafe.

Ontogénesis de la estratificación moral.

La perspectiva filogenética ha de complementarse con una *perspectiva ontogenética*. Una de las aproximaciones más interesantes al respecto es la del psicólogo Marc Hauser. Este plantea la existencia de una serie de estructuras morales que funcionan, en analogía con el lenguaje, como una gramática moral. Estas estructuras tendrían un carácter universal, funcionando como un marco para la adquisición de ciertos contenidos, sin determinar cuáles son estos, que vendrán marcados por el entorno cultural en que el individuo se desarrolle. Recordemos que la contingencia en los contenidos no significa que exista un relativismo moral, a pesar de esta contingencia la filosofía moral investiga cuáles son los criterios que permiten fundamentar ciertos contenidos, desechando otros como éticamente válidos.

En *La mente moral: Cómo la naturaleza ha desarrollado nuestro sentido del bien y del mal* (2008) Hauser expone ampliamente esta tesis y los argumentos que la apoyan. Su propuesta toma como modelo la que Chomsky hizo en relación al lenguaje (Chomsky, 1987), y la plantea como análoga en muchos aspectos. La idea de una gramática moral universal guarda un paralelo evidente con la idea de una gramática lingüística universal. En ambos casos se plantea la existencia de un equipamiento psicobiológico dado de forma innata y que predispone al individuo para la adquisición de ciertos contenidos en el seno

de una cultura. Estos contenidos se corresponden con los diferentes sistemas morales y las distintas lenguas, respectivamente.

La gramática universal, que se encuentra en el corazón de nuestra facultad lingüística y forma parte de las dotes innatas de nuestra especie, ofrece toda una serie de herramientas para construir lenguajes concretos. Una vez que hemos adquirido nuestra lengua vernácula, hablamos y comprendemos lo que otros dicen sin necesidad de razonar y sin tener conciencia de las reglas y principios subyacentes. Sostengo que nuestra facultad moral está equipada con una *gramática moral universal*, una caja de herramientas apta para construir sistemas morales concretos. Una vez que hemos adquirido las normas morales propias de nuestra cultura (proceso que se parece más al crecimiento de un miembro corporal que a la asistencia a la escuela dominical para recibir instrucción sobre vicios y virtudes), juzgamos si determinadas acciones son permisibles, obligatorias o están prohibidas, sin necesidad de razonar conscientemente y sin tener acceso explícito a los principios subyacentes. (Hauser, 2008: 17-18)

Lo que Hauser plantea es un conjunto de predisposiciones biológicas que establecen las condiciones en las cuales cualquier sistema moral puede ser adquirido y desarrollado. Estas estructuras psicológicas y neuronales innatas no especifican qué contenidos se alojarán en ellas, sino el modo en que estos contenidos van a ser adquiridos. Está hablando, en definitiva, de la moral como estructura psicobiológica. Esta gramática sirve como guía, como límite y como recipiente de los diversos contenidos morales.

Decir que estamos equipados con una gramática moral universal es decir que hemos desarrollado principios generales pero abstractos para decidir qué acciones están prohibidas, son lícitas u obligatorias. Estos principios carecen de contenido específico. No hay principios que dicten qué actos sexuales, altruistas o violentos son permisibles. Nada en nuestro genoma codifica si el infanticidio, el incesto, la eutanasia o la cooperación son lícitos y, si lo son, con qué individuos. Y la manera más sencilla de ver que esto debe ser así es darse cuenta de que cada niño, según cuales sean sus orígenes culturales, adquirirá un sistema moral distinto. La gramática moral universal es una teoría acerca de los principios que permiten a los niños construir una gama amplia, pero finita, de sistemas morales distintos. (Hauser, 2008: 486-487)

Una exposición detallada del funcionamiento de esta gramática moral universal ofrecerá no una explicación de los factores conscientes que intervienen en el juicio moral, sino una explicación de las estructuras y mecanismos subyacentes, que no suelen ser conocidos por el individuo que emite el juicio.

El resultado es una explicación profusamente detallada de cómo una gramática moral inconsciente y universal subyace a nuestros juicios acerca de lo que está bien y lo que está mal. (Hauser, 2008: 18)

Tanto el usuario de un lenguaje como el de un sistema moral puede hacer uso eficiente de estos sistemas sin necesidad de conocer los mecanismos psicobiológicos subyacentes a ellos. La tarea del psicólogo moral y del neurocientífico consistirá en revelar el funcionamiento de las estructuras que subyacen a cualquier sistema moral.

En lingüística es importante diferenciar gramáticas concretas o individuales, referidas a las reglas que componen cada lengua, y gramática universal, concerniente a las reglas y principios que rigen el aprendizaje de cualquier lengua y que vienen dados de forma innata como una estructura psicobiológica (Hauser, 2008: 66). De forma análoga podemos diferenciar los sistemas morales concretos de la gramática universal de la que depende su adquisición.

Cada sistema, lingüístico o moral, puede estar gobernado por una serie de principios y normas internos que lo dotan de sentido, y que pueden variar, siendo contingentes e históricos. Estas variaciones quedan sujetas a la gramática moral universal, que impone las posibilidades de adquisición y desarrollo de contenidos.

Si hay una gramática moral universal, los principios son fijos, pero la gama potencial de sistemas morales no lo es. Esa gama potencial está solo limitada por las posibilidades lógicas del cerebro y un cierto grado de inercia histórica. (Hauser, 2008: 486)

Los conceptos de moral como estructura y moral como contenido son clave para comprender la propuesta de Hauser. Una gramática moral universal daría cuenta de las estructuras psicobiológicas que orientan la conducta probabilísticamente sin determinarla, dejando margen para que diferentes respuestas sean posibles ante un mismo estímulo, permitiendo la adquisición de diferentes sistemas morales en el seno de diferentes culturas. Dicho de otro modo, Hauser describe las estructuras psicobiológicas de la moralidad, que dotarían al ser humano de una gramática universal, entendida como predisposición biológica estructural para la adquisición de contenidos morales, adquiridos culturalmente.

Cabe añadir que esta adquisición de contenidos no es un proceso de escritura de contenidos dentro de los márgenes que establece esa gramática universal, es más bien un

proceso sinérgico, en el que las estructuras condicionan la adquisición de contenidos mientras que la adquisición de estos va modificando a su vez las propias estructuras. Dado que el desarrollo cerebral se produce mayoritariamente en la interacción del sujeto con el medio (Evers, 2010; Cortina, 2012a; Jácomo y Serrão, 2013), la gramática universal no sería algo estático, sino que contaría con cierta flexibilidad en la que influirían factores ontogenéticos y epigenéticos. Aranguren refiere a esta cuestión al hablar de una segunda naturaleza por apropiación. La distinción entre moral como estructura y moral como contenido, advierte, no presenta límites abruptos en la realidad, ya que los contenidos morales afectan a la propia estructura, modificándola en algún sentido.

La moral consiste no solo en el ir haciendo mi vida, sino también —y esta es la vertiente que ahora estamos examinando— en la vida tal y como queda hecha: en la incorporación o apropiación de las posibilidades realizadas. La moral resulta ser así algo *físicamente real* o, como decía Aristóteles, una *segunda naturaleza*. (Aranguren, 1994: 56)

Ahora bien: esta apropiación real de posibilidades, buenas o malas, va conformando en una «segunda naturaleza» mi personalidad. O, como decía Zubiri, al apropiarme de mis posibilidades constituyo con ello mi habitud en orden a mi autodefinición, a la definición de mi personalidad. Sobre mi «realidad por naturaleza» se va montando una «realidad por apropiación», una «realidad por segunda naturaleza» que, inseparablemente unida a aquella, la conforma y cualifica según un sentido moral. Mi realidad natural es mi propia realidad, en tanto que recibida; mi realidad moral es mi propia realidad, en tanto que apropiada. Porque al realizar cada uno de mis actos voy realizando en mí mismo mi *êthos*, carácter o personalidad moral. (Aranguren, 1994: 56-57)

Esta interrelación entre la estructura moral y los contenidos que en ella se alojan es, por tanto, no una mera relación de continente y contenido, sino que además el contenido modifica la realidad sobre la que se aloja. El efecto realizativo de los contenidos morales sobre las propias estructuras naturales remite al concepto de epigénesis, que en el ámbito de la psicología del desarrollo es descrita como un proceso de interacción entre factores genéticos y no genéticos (gen-ambiente) (Piaget). Puesto que aproximadamente un 70% del desarrollo del cerebro se produce después del alumbramiento (Evers, 2010), en la interacción del sujeto con el medio a lo largo de su vida, no es difícil comprender y aceptar la propuesta de Aranguren desde una perspectiva psicológica y neurocientífica.

La asimilación reduccionista entre estructuras y contenidos.

La tercera tarea que nos proponemos llevar a cabo tiene una relevancia especial en relación a la cuestión de la fundamentación ética, esta tarea consiste en denunciar el

reduccionismo cometido por algunos neurocientíficos, que sostienen que partiendo de sus investigaciones científicas pueden extraerse conclusiones normativas. En lo que concierne a la moral como estructura y como contenido el principal error de reduccionismo científico es el cometido al considerar que las estructuras morales, compuestas por la realidad neuronal, psicológica, genética y biológica de nuestra especie, ofrecen información suficiente para resolver el problema de qué códigos éticos deben gobernar nuestras acciones en sentido moralmente normativo. Dos de los autores que incurren en este error son Gazzaniga y Mora, a los que atenderemos a continuación. El propio Aranguren, al plantear el estrato estructural y el de contenidos de la moral, nos advierte del riesgo del reduccionismo. Nos habla de la conveniente posibilidad que plantea Le Senne de abrir “desde la psicología, el ámbito de la moral, sin caer, como los psicólogos, en la confusión de ambas esferas” (Aranguren, 1994: 52).

La ética se halla anclada en la psicología y aún en la biología, lo que no significa, es claro, que «toda moral sea de esencia biológica», como opinó Bergson, ni aun tomando este último término, como él lo hacía, en el más amplio y contundente sentido. (Aranguren, 1994: 53)

Aranguren reconoce acertadamente que el papel de la psicología y de la biología es clave para comprender el hecho de lo moral, sin que ello signifique que todo lo relativo a este hecho sea explicable desde una perspectiva científica naturalista.

Gazzaniga: una ética integrada en el cerebro.

En 2005 Gazzaniga publica *El cerebro ético*, un libro en el cual se adentra en cuestiones de neuroética, disciplina a la que define como emergente, motivo por el cual el autor trata de esbozar algunas de las cuestiones más relevantes en el campo. Nos interesa especialmente la cuestión de si es posible descubrir en el cerebro un código ético que rijan nuestra conducta en sentido moral y normativo. El autor afirma que sí es posible hacer tal cosa desde las neurociencias.

En suma, me gustaría defender la idea de que podría existir un conjunto universal de respuestas biológicas a los dilemas morales, una suerte de ética integrada en el cerebro. Espero que descubramos pronto esos principios éticos, los identifiquemos y empecemos a vivir en consonancia con ellos. Creo que ahora ya vivimos así de un modo fundamentalmente inconsciente, pero sería posible eliminar mucho sufrimiento, guerras y conflictos si conviniésemos en vivir de acuerdo con tales principios de manera consciente. (Gazzaniga, 2006: 17-18)

Gazzaniga plantea la existencia de un conjunto de predisposiciones biológicamente configuradas que sería universal en tanto que se darían de forma innata en la especie humana. Sugiere que este conjunto de predisposiciones puede ser entendido como una ética que viene integrada en el cerebro, heredada genéticamente y adquirida en el proceso evolutivo. Podría pensarse que quizás lo que plantea el autor es una suerte de estructura biológica predispositiva, en el sentido en que Hauser plantea la gramática moral universal, pero esta posibilidad queda disuelta en el momento en que Gazzaniga afirma que sería conveniente vivir de forma consciente de acuerdo a esos principios éticos forjados por la evolución en nuestro cerebro. Es cierto que existen pautas de comportamiento que condicionan nuestra conducta desde nuestra biología, y que estas pautas operan muy a menudo de forma inconsciente —Gazzaniga cita a Hauser y a Greene entre otros—, ahora bien, ¿deben regir estas pautas nuestra conducta en un sentido ético?

Pretender que las pautas comportamentales que resultan biológicamente provechosas sean las más éticas es problemático en varios sentidos. En primer lugar, dar el salto deductivo de un hecho a una norma moral es falaz. Que por naturaleza existan en nuestro cerebro predisposiciones a actuar según unas pautas de comportamiento determinadas no significa que debamos de actuar según tales pautas en un sentido ético y normativo. En segundo lugar, podría argumentarse que podría darse el caso de que lo que aconseja la reflexión ética y lo que predispone la naturaleza coincida, en tal caso nuestra biología nos predispondría para aquello que además es justificable moralmente. Respecto a ello hay dos cosas que objetar. La primera es que esa coincidencia sería casual, no causal. La segunda es que tal caso no sucede, las predisposiciones conductuales que nuestra biología predispone son egocéntricas y endogámicas, están dirigidas a que el individuo se beneficie a sí mismo y al grupo que le es más cercano y familiar. Por su parte, la ética, al menos en la mayoría de sus propuestas, plantea el establecimiento de normas imparciales y de corte universal. La disonancia entre los dictados de la biología y la ética es muy marcada. Si bien es cierto que existen algunos casos en los que ciertas tendencias biológicas pueden coincidir con normas morales éticamente aceptadas, este tipo de coincidencias no se deben a que la biología fundamente la norma en sentido moral. Por tanto, biología y ética difieren significativamente respecto a los códigos de conducta que cada una de ellas plantea, aún si existen nexos entre una y otra, marcados, entre otras cosas, por nuestras emociones de prosocialidad (Cortina, 2012a: 59-61; Cortina, 2017).

En el Capítulo 10 de *El cerebro ético* el autor reflexiona sobre esa ética universal grabada en el cerebro, sería de esperar que este capítulo plantee una propuesta sólida, aún si contiene errores. Sin embargo, la propuesta de Gazzaniga no ofrece contenidos para esa ética universal, siendo deficitaria en esta cuestión, que debería de componer su núcleo. Únicamente plantea la existencia de unas estructuras morales dadas por la biología, tesis que compartimos, y la necesidad de descubrir qué códigos éticos se derivarían de esas estructuras, siendo estos los códigos que deben regir nuestra conducta en sentido moralmente normativo, tesis respecto a la cual disintimos. La inconcreción de la propuesta hace que esta sea deficiente, pues en realidad tan solo se ha planteado de forma borrosa la posibilidad de una ética universal fundamentada en la biología, sin dar argumentos suficientes a su favor ni concretar los aspectos, principios y normas que compondrían esta ética. El capítulo se cierra dejando abierta esta cuestión, lo cual denota la vaguedad de los argumentos del autor a favor de la tesis que pretende defender.

Creo, por lo tanto, que no debemos buscar una ética universal que abarque verdades absolutas, sino la ética universal que nace del hecho de ser humano, que es claramente contextual, sensible a la emoción y orientada al refuerzo de la supervivencia. Por eso es tan difícil formular reglas absolutas para la vida en las que coincidamos todos. Pero la constatación de que la moralidad es contextual y social, y de que se basa en mecanismos neuronales, puede ayudarnos a determinar ciertas maneras de tratar cuestiones éticas. Ése es el imperativo de la neuroética: partir de una constatación científica —la observación de que el cerebro reacciona ante las cosas según su configuración— para contextualizar y discutir los instintos viscerales que aportan mayores beneficios —o las soluciones más lógicas— en determinados contextos. (Gazzaniga, 2006: 179)

La lectura de este fragmento parece sugerir que Gazzaniga plantea la necesidad de comprender la ética teniendo en cuenta la base biológica que subyace a esta, sin que necesariamente determine qué contenidos éticos son los más adecuados. Al hablar de la influencia del contexto y de la búsqueda de las soluciones más lógicas parece que el autor contempla la posibilidad de que los contenidos puedan variar según los argumentos aportados a su favor y el contexto social y cultural en que se planteen. Tal postura sería adecuada, sin embargo no es la que propone realmente. Al contrastar este fragmento con el que hemos citado anteriormente vemos que su propuesta es otra. Lo que Gazzaniga plantea es que nuestra biología y nuestro cerebro ofrecen unos códigos que permiten un desarrollo en el medio social y ofrecen soluciones beneficiosas y lógicas para el máximo beneficio biológico en determinados contextos, sugiriendo además que esos códigos deben regir nuestra conducta en sentido moral. Sin embargo, no especifica qué códigos

son estos. Todo ello constituye su propuesta como deficitaria y problemática en los aspectos señalados.

Mora: hacia una ética universal de base cerebral.

Una perspectiva muy cercana a la anterior es la de Francisco Mora, quien, en *Neurocultura: Una cultura basada en el cerebro* (2007), atiende a la relación entre neurociencias y ética en un capítulo titulado “Neuroética”. Sugiere que a partir de los códigos de conducta grabados en el cerebro pueden extraerse códigos de conducta con validez normativa en sentido moral. Una aproximación a esta tesis podemos verla en el siguiente fragmento, si bien en él no aclara del todo la cuestión.

En sus raíces, la neuroética tiene su pilar básico, más robusto, en la aceptación de que lo que hoy llamamos ética depende, en toda su dimensión, del funcionamiento del cerebro y, en particular, de ciertos sistemas cerebrales trabajando en un contexto social. (Mora, 2007: 72)

Mora afirma que la ética depende en toda su dimensión del funcionamiento del cerebro. Es cierto que, en tanto que resultado de una actividad cerebral, la ética, como cualquier otra actividad humana, depende enteramente del cerebro. Lo que Mora no aclara es si esa dependencia se da en tanto que no es posible hacer ética sin un cerebro, o si, por el contrario, lo que quiere decir es que todo lo que propone la ética —contenidos morales, condiciones de justificaciones del juicio y la acción, propuestas de fundamentación, etc.— viene dado por la biología, siendo éticos en virtud de que esta marca los códigos normativos que han de regir nuestras vidas moralmente. La primera de las opciones es aceptable, la segunda no. Debemos seguir indagando en el texto para intentar esclarecer la postura del autor. La afirmación más comprometida con esta idea la hace al final de la obra, donde plantea que el conocimiento del cerebro puede ofrecer herramientas de corte más universal, como un lenguaje y una ética compartidos universalmente.

Y también una ética universal a través de la cual se puedan alcanzar valores y normas morales asumidos y respetados por todos los seres humanos. Porque la ética, ya no concebida como emanada de Dios (diferente para cada grupo social o geográfico), o grupos religiosos sin Dios, sino anclada en las raíces evolutivas del hombre y su cerebro, puede encontrar ahora un punto único y común de anclaje. La ética tiene claramente un origen biológico y aunque sea, como lo es, un producto en parte evolutivo, es auténtica y genuinamente humana. Los valores éticos tan diferentes, para grupos étnicos tan diferentes, pueden converger en reglas y normas establecidas por la neuroética basadas en el funcionamiento del cerebro humano, base común a todos los

hombres. Ya lo decía Churchland: «Los mejores candidatos actuales para entender el origen de la moralidad son naturalistas». (Mora, 2007: 159)

Este fragmento disuelve todas las dudas, mientras que a lo largo de la obra Mora no se ha pronunciado de forma explícita respecto a la cuestión, en el Epílogo, donde encontramos el fragmento anterior, se aventura a hacerlo, planteando la idea como uno de los corolarios de todo su texto. La ética universal que plantea sería capaz de proponer valores y normas morales que serían asumidos y respetados por todos los miembros de la especie, en tanto que todos comparten un cerebro humano. Esto significaría que el cerebro no es solo una estructura que posibilita la agencia moral, entendida como facultad que permite la moralidad, siendo contingentes los contenidos que el agente pueda tener, sino que el cerebro ofrece criterios para establecer valores y normas con contenido, que serían universalizables en tanto que toda la especie comparte una estructura que otorga una predisposición a estas normas y valores, el cerebro. Esta idea es la que refiere Mora al decir que este órgano es el punto de anclaje único y común para una ética compartida por todo ser humano. Por todo ello su propuesta, como la de Gazzaniga, es también problemática y susceptible de crítica.

La necesidad de distinguir entre bases y fundamentos.

Hay una confusión latente bajo los planteamientos de Gazzaniga y de Mora. Si bien coincido con este último, y con Churchland (2013), en que el naturalismo es el mejor candidato para comprender el origen de la moralidad, esto no significa que también lo sea para comprender la ética en todas sus dimensiones (Cortina, 2012a; Richart, 2016; Ortega Esquembre, 2016). Moralidad y ética son cosas diferentes. La moralidad puede comprenderse como un hecho natural, los seres humanos son seres morales. Qué hace posible esa moralidad es algo que las ciencias naturales pueden tratar de explicar, en tanto que las capacidades cognitivas que producen la facultad de la agencia moral son susceptibles de ser estudiadas empíricamente. Ahora bien, aun si la ética es el resultado del ejercicio de esas capacidades, esto no implica que sea reducible a ellas. Qué procedimientos permiten establecer y fundamentar normas y valores morales, qué entendemos por justicia, qué entendemos por responsabilidad, etc., no son aspectos de la ética que estén determinados por nuestra biología, a pesar de que la reflexión ética que los aborde esté posibilitada por las condiciones que esta presenta en el ser humano. La

labor de la ética no es reductible, por tanto, a la comprensión de la moralidad en tanto que hecho psicobiológico y natural.

El problema de este tipo de propuestas es que incurren en reduccionismo naturalista al identificar las estructuras que permiten lo moral y lo ético con el propio contenido de la ética. Cortina hace una aportación esclarecedora al distinguir los conceptos de *base* y de *fundamento* en relación al cerebro y lo moral. Según expone Cortina, hablar de *bases cerebrales* de la moral significaría hacer referencia a las estructuras psicobiológicas constituidas en el cerebro que permiten la existencia de lo moral, mientras que hablar de *fundamentos cerebrales* supondría dar a esas estructuras el carácter de razón última para la justificación de las normas que han de regir nuestra conducta en sentido ético y normativo.

[...] no es lo mismo hablar de una *base cerebral* de la conducta moral, como hay una base fisiológica, psicológica y social, que intentar *fundamentar* la conducta moral en elementos cerebrales. (Cortina, 2012a: 94)

La fundamentación de la ética no depende de las estructuras cerebrales, sino de las razones que se pueden aportar para sustentar unas normas en detrimento de otras. Las condiciones de la fundamentación ética son aquello sobre lo que versa la reflexión filosófica de las diferentes teorías éticas. Si bien esta reflexión participa de ciertas estructuras psicobiológicas, gracias a las cuales es posible, no es, sin embargo, reductible a ellas. La *fundamentación* de la ética se da en un nivel de reflexión que tiene una estructura de sentido diferente a la propia de la biología, aunque dependa de esta como condición de posibilidad, siendo su *base*.

Por eso es importante distinguir entre las *bases* de una ética universal, que serán por supuesto cerebrales, pero también mentales y sociales, y el *fundamento* de una ética universal, que nos permite dar razones morales ante la pregunta por el carácter exigitivo de normas, valores, sentimientos y virtudes a los que llamamos morales. (Cortina, 2012a: 96)

En virtud de ello hay que adoptar una actitud crítica ante aquellas posturas que confunden las *bases* cerebrales de la moral con el *fundamento* de la ética. Una pretensión de este tipo resulta reduccionista y comete errores lógicos falaces, pretendiendo que datos científicos sobre nuestra biología y nuestro cerebro impliquen conclusiones normativas sobre cómo debemos actuar en sentido ético.

4.2. Libertad como estructura moral: qué puede decir la neurociencia sobre el libre albedrío

La cuestión de la libertad y el determinismo en la conducta humana ha suscitado un debate de forma recurrente a lo largo de la historia del pensamiento, podemos encontrarlo desde Heráclito hasta corrientes más actuales del pensamiento como el materialismo histórico, el determinismo genético, el económico, y ahora el neurocientífico, por mencionar algunos de ellos. La neurociencia pone sobre el tapete algunas cuestiones en torno a la libertad, entendida como libre albedrío. La principal pregunta que se plantea al respecto es si existe la voluntad libre, si somos libres a la hora de actuar, o si nos vemos irrevocablemente conducidos a actuar de forma determinista por acción de nuestras neuronas. Si la respuesta es que no hay libertad, conceptos como el de responsabilidad o imputabilidad perderían todo su sentido, y se vendría abajo el edificio de la moral, el derecho, la política, la religión, la economía y la ciencia, que se articulan en gran medida en torno a esa idea de libertad que se había dado por supuesta. Además, la admiración por alguien, la culpabilidad o el agradecimiento perderían todo sentido si el admirado, el culpable o el benefactor no es en realidad responsable de lo que ha hecho. A ello hemos de sumarle la riqueza de posibilidades de la vida política, moral y económica, que se plantea a veces no como resultado de una naturaleza humana, sino como un artificio o construcción, de lo cual la libertad podría dar buena cuenta. La ausencia de libertad, por estos motivos, supondría una pérdida de sentido para la moral, la política y el derecho tal y como los conocemos, especialmente en lo concerniente a la responsabilidad moral y a la legal, y coartaría la posibilidad de defender, en el sentido en que históricamente se ha hecho, las libertades que plantean los derechos civiles y políticos. Por todo ello hemos de analizar qué posibilidades y límites encuentran las neurociencias a la hora de investigar sobre la cuestión del libre albedrío y del determinismo. (Cortina, 2012a: 151-213)

Son muchos los problemas metodológicos y conceptuales a los que se enfrentan los investigadores al tratar de resolver científicamente la pregunta por la existencia del libre albedrío. Nuestro objetivo será someter a crítica uno de los presupuestos de estas investigaciones, la posibilidad de cuantificar la libertad como un hecho natural. Para ello expondremos en primer lugar las dificultades relativas a esta pretensión. Tras ello estableceremos cuáles son las condiciones de posibilidad de la libertad en aras de someter

a crítica las perspectivas psicológica y neurocientífica. Argumentaremos que ambas son incapaces de ofrecer conocimiento sobre la existencia de tales condiciones.

Con el auge de las neurociencias ha nacido la pretensión de investigar de forma naturalista, mediante el estudio del cerebro, si existe el libre albedrío en la conducta humana o si por el contrario esta está determinada de forma unívoca. La libertad es una de las condiciones necesarias para que una acción sea genuinamente moral (Ayala, 2006). La posibilidad de actuar de diferentes formas y de poder escoger entre ellas es condición indispensable para la moralidad, si la conducta estuviese determinada por una cadena causal en la que se imbricasen nuestra biología, nuestro entorno y todo el devenir del universo, en un sentido más amplio y holístico, entonces formaríamos parte de una complejísima red en la que unos hechos conducen a otros de forma unívoca, incluyendo nuestra conducta. El sentimiento de libertad, en tal caso, no sería más que una ilusión. Si se demostrase que esto es así, conceptos como el de responsabilidad y el de elección, y toda nuestra concepción no solo de la moral y la ética, sino de la realidad humana, del derecho, de la sociología, de la historia, etc., se verían comprometidos, y habrían de ser revisados. Sin embargo, si se demostrase que existe el libre albedrío, podríamos desembarazarnos de estas preocupaciones y continuar con nuestra concepción del ser humano como un ser libre (Cortina, 2012a, 151-153; Cortina, 2012b; Farah, 2012; Illes y Sahakian, 2012; Kant, 1992, 2000; Roskies, 2010; Pallarés-Domínguez y González Esteban, 2016).

Considero que existe una imposibilidad en las ciencias empíricas, y más concretamente en las neurociencias, para resolver la cuestión de si el ser humano puede actuar de forma libre o si se halla determinado de forma unívoca en su conducta. Para respaldar esta tesis pretendo poner en tela de juicio las pretensiones neurocientíficas que defienden la posibilidad de cuantificar la libertad como un hecho natural. Para ello, en primer lugar, expondré los principales problemas y cuestiones relativos a la libertad y su cuantificación desde las ciencias empíricas. En segundo lugar, atendiendo a Evers (2010) y Cortina (2012a), estableceré cuáles son las condiciones de posibilidad de la libertad, en aras de aclarar si es posible descubrir la existencia de tales condiciones desde la perspectiva naturalista de las ciencias empíricas y, concretamente, de las neurociencias. Tras ello atenderemos a Aranguren, para poner de relieve cómo una perspectiva meramente psicológica de la libertad no puede ofrecer conclusiones acerca de si la

conducta se configura unívocamente o si, por el contrario, permite diversas acciones ante un mismo estímulo. Seguidamente someteremos a crítica la investigación naturalista de las neurociencias en torno a la cuestión de la libertad, centrando el análisis en el experimento de Libet (1999), que tomaremos como eje de la reflexión, prestando especial atención a la crítica que Cortina vierte sobre este. Atenderemos también a Habermas (2006) en este punto, cuyas críticas al experimento quedan recogidas por Cortina. Finalmente se plantearán las conclusiones derivadas de todo lo expuesto.

¿Puede cuantificarse la libertad empíricamente?

La existencia de sujetos libres precisa que estos hayan alcanzado cierto umbral de desarrollado y complejidad cognitiva²⁴. Es preciso que se dé la existencia de estructuras cerebrales capaces bien de romper las cadenas causales de nuestra biología o bien de generar una contingencia de consecuencias ante una misma causa (Evers, 2010: 86-97), siendo además el sujeto quien puede escoger por autodeterminación una de las opciones disponibles (Evers, 2010: 86-97; Cortina, 2012a: 161). Sobre las condiciones de posibilidad de la libertad volveremos más adelante. Ahora bien, no sabemos si la evolución ha forjado seres realmente libres, o si más bien ha generado seres que se autoperciben como libres²⁵ (Farah, 2012), es precisamente esta cuestión la que las neurociencias pretenden resolver.

Ya que la existencia o inexistencia del libre albedrío presenta como base las estructuras cerebrales que configuran la decisión y la acción, es comprensible que hayan aparecido tentativas neurocientíficas de cuantificar la libertad o su inexistencia como hecho natural y neurobiológico. Ahora bien, nada garantiza que la ciencia empírica sea capaz de ofrecer resultados significativos al respecto. La dificultad de llevar a cabo una investigación sobre la existencia del libre albedrío desde una perspectiva naturalista radica en determinar si es o no posible la cuantificación de la libertad o de su inexistencia como un hecho natural, de serlo habrá que aclarar qué consideraciones metodológicas

²⁴ Atendiendo a Ayala, si bien este no plantea esta idea explícitamente, su propuesta en torno a las condiciones de posibilidad de una conducta ética sugiere que el nivel de desarrollo cognitivo está asociado a la posibilidad de emergencia de tales conductas, siendo una de sus condiciones de posibilidad el libre albedrío. (Ayala, 2006; Richart, 2016)

²⁵ Esta autopercepción ilusoria podría deber su origen filogenético bien por ofrecer una ventaja evolutiva a la hora de regular la conducta, o bien gracias a ser una consecuencia no prevista por la evolución (*by-product*) resultante de la concurrencia de ciertas facultades mentales, que permitiría usos no previstos por la evolución de estructuras y capacidades cerebrales desarrolladas para otras funciones. (Gould y Lewontin, 1979; Gould y Vrba, 1982; Richart, 2016).

habrían de ser tenidas en cuenta. La neurociencia plantea, con el experimento de Libet como uno de sus estandartes, siendo a su vez muy problemático y susceptible de crítica, la posibilidad de desvelar la respuesta a la cuestión de si estamos determinados o somos libres. ¿Pero está realmente la neurociencia capacitada para ofrecer una respuesta a una cuestión de este tipo? ¿Puede ser la existencia o la inexistencia del libre albedrío cuantificada como un hecho natural observable?

A lo largo del epígrafe argumentaré a favor de la siguiente tesis: si bien la existencia o inexistencia del libre albedrío en el ser humano ha de producirse como un hecho natural, existen dificultades, quizás una imposibilidad, para cuantificar este hecho desde el método científico, lo cual otorga un papel preminente a la filosofía a la hora de abordar la cuestión.

Los experimentos llevados a cabo hasta la fecha no son capaces de ofrecer conclusiones claras al respecto. A pesar de que el método científico tiene entre sus competencias explicar y describir la realidad –así como promover y diseñar la aplicación técnica y tecnológica que permite ese conocimiento– eso no significa necesariamente que la ciencia sea capaz de conocer toda la realidad natural existente. Simplemente algunos aspectos de lo real escapan por completo a los sistemas de medición con los que la ciencia cuenta. Es posible que la investigación en torno a la libertad se enfrente a una realidad de este tipo. Es decir, se ha de considerar la posibilidad de que la mera cuantificación de la actividad y estructuras cerebrales, que subyacen bajo las conductas susceptibles de ser libres, sean constitutivamente insuficientes para desvelar si de hecho podemos obrar haciendo uso del libre albedrío.

Las condiciones de posibilidad de la libertad.

Antes de someter a análisis y crítica la posibilidad de cuantificar el libre albedrío desde las ciencias empíricas es preciso establecer cuáles son las condiciones de posibilidad de una conducta realmente libre. Una vez establecidas estas condiciones estaremos en posición de analizar si la perspectiva naturalista de las ciencias empíricas es capaz de resolver la cuestión desde su método. Recurriremos a Evers y Cortina para señalar cuáles son las condiciones de posibilidad necesarias y suficientes para que una conducta pueda ser considerada como libre.

Kathinka Evers en *Neuroética: Cuando la materia se despierta* (2010) aborda la cuestión de la libertad en un capítulo titulado “El cerebro responsable, el libre albedrío y la responsabilidad personal a la luz de las neurociencias” (2010: 74-112). A lo largo del capítulo Evers expone los términos generales del debate libertad-determinismo para posicionarse en ese debate y exponer su postura respecto a la libertad y la responsabilidad en el ámbito de las neurociencias. Sostiene que el libre albedrío es una estructura neuronal fundamental, junto a otras como el espacio, el tiempo o la causalidad. Con independencia de si la experiencia de libertad es una mera ilusión, Evers sostiene que podemos hablar de actos libres en el sentido de voluntarios, aunque fuesen una construcción causalmente determinada por procesos neuronales no conscientes. Además, suscribe la idea de que puede existir una *causalidad contingente* compatible con la libertad de elección, según la cual ante una misma causa pueden darse dos o más consecuencias diferentes, entre las cuales la elección sería posible. El teorema de la variabilidad en neurociencias supone un apoyo empírico para tesis de este tipo (Evers, 2010: 86-97). Asimismo, considera que se puede ser responsable a pesar de los procesos no conscientes que subyacen a las decisiones y a las acciones, ya que somos personalmente responsables de esos estados y procesos en la medida en que podemos influir en ellos. A lo largo de su reflexión ofrece dos condiciones que han de darse para que un sujeto sea libre:

- a) En primer lugar, a la hora de actuar, el sujeto ha de poder producir al menos dos conductas diferentes ante una misma situación, o dicho de otro modo, ha de poder generar al menos dos consecuencias diferentes ante una misma causa. Es lo que Evers considera como una *contingencia causal*, pueden producirse dos mundos posibles a partir de una causa o conjunto de causas, de tal modo que aun sin escapar de la cadena de causalidad existe la posibilidad de bifurcación dentro de la misma.
- b) Esa resolución en la contingencia de la acción no debe producirse al azar ni por indeterminación, sino de forma autodeterminada. La libertad es autodeterminación, el sujeto ha de poder determinar qué acción llevar a cabo ante la opción de ejecutar varios cursos de acción posibles. Entran en juego aquí la reflexión, las razones y la decisión consciente, qué papel juega cada uno de estos elementos es planteado de diferente modo según diferentes autores.

Esas dos condiciones no se producen solo en el ámbito de lo mental, dado que lo mental arraiga en las estructuras biológicas y cerebrales del sujeto dependerá de su arquitectura cerebral la posibilidad de generar un espacio de decisión entre al menos dos posibles cursos de acción, haciendo que además esa decisión se produzca de forma autodeterminada.

Cortina se ocupa también de la cuestión de la libertad. Esta, indica, suele entenderse como una propiedad de la voluntad, y precisa de tres condiciones (Cortina, 2012a: 155-156):

1. Entre un conjunto de posibilidades la que elegimos está en nuestras manos.
2. La fuente de nuestras acciones está en nosotros, no en algo sobre lo que no tenemos control.
3. Podríamos haber actuado de manera diferente, había alternativas y la posibilidad real de optar por ellas.

La condición *a* de Evers aglutina las condiciones 1 y 3 de Cortina, mientras que la condición *b* corresponde con la 2, respectivamente. Podemos por tanto identificar una coincidencia en la postura de ambas pensadoras.²⁶

Las condiciones de posibilidad de una conducta libre —a las que referiremos como *L1* y *L2* durante la exposición— son, por tanto, las siguientes:

L1) dadas unas condiciones ha de existir la posibilidad real de escoger entre dos cursos de acción diferentes.

L2) la elección se ha de dar por autodeterminación, no deterministamente, por azar o por indeterminación.

Estas condiciones a su vez implican una serie de presupuestos que toda investigación científica debería de resolver. Desglosar estas dos condiciones en los siguientes presupuestos consiste, en realidad, en confeccionar una lista más específica de

²⁶ Añade Cortina que en caso de que la libertad no existiese y que científicamente pudiese demostrarse, hay quien plantea un *compatibilismo* entre ese conocimiento y el vivir como si la libertad existiese, aduciendo que es como mejor sabemos organizar nuestra vida. ¿Pero acaso no sería esto ignorar un conocimiento al que aspiramos para mejorar nuestras vidas? Si la libertad no existiese algunas disciplinas como el derecho deberían ser reestructuradas de forma profunda, afirma la autora.

condiciones de posibilidad, que aparecen planteadas por *L1* y *L2* tan solo de forma implícita.

Entre los presupuestos de *L1* —referidos como *L1a* y *L1b*— se encuentran los siguientes:

L1a) Existe bien una fisura en la cadena causal, o bien una bifurcación posible en la misma (contingencia causal);

L1b) existe la posibilidad, brindada por nuestro cerebro, bien de generar una causa que produzca una decisión y una acción (el *yo* sería esta causa, ver *L2b*), bien de dirigir la conducta hacia la opción escogida en una contingencia causal, respectivamente.

Por su parte entre los presupuestos de *L2* —referidos como *L2a* y *L2b*— encontramos los siguientes:

L2a) Existe un *yo*, que sería el decisor y el actor;

L2b) el *yo* tiene la capacidad de operar como causa en la decisión y en la acción, bien como una nueva causa dada una fisura en la cadena causal, bien como causa de la elección entre las opciones posibles en una contingencia causal.

Así las dos condiciones de posibilidad del libre albedrío expuestas por Evers y Cortina (*L1* y *L2*) revelan un total de cuatro condiciones de posibilidad (*L1a*, *L1b*, *L2a* y *L2b*), que quedan especificadas al desglosar los presupuestos de aquellas dos.

Una investigación naturalista que pretenda argumentar a favor de la existencia del libre albedrío tiene que poder demostrar que todas estas condiciones se dan, para demostrar que el libre albedrío no existe se tendrá que probar que al menos una de estas condiciones no se da. Serán insuficientes todas aquellas propuestas que pretendan argumentar a favor del determinismo simplemente por haber sido incapaces de demostrar la existencia de la libertad, a tales propuestas se les ha de exigir que demuestren su inexistencia, si es que es posible hacer tal cosa. De igual modo serán insuficientes aquellas propuestas que se posicionen a favor de la existencia del libre albedrío sin haber logrado demostrar que se cumplan todas las condiciones de posibilidad de su existencia.

En los próximos apartados atenderemos a dos ejemplos de propuestas que cometen el error señalado, no ofreciendo una demostración del cumplimiento o incumplimiento de las condiciones de posibilidad del libre albedrío al afirmar su existencia o su inexistencia, respectivamente. El primer caso corresponde a la perspectiva psicológica ofrecida por Aranguren en su *Ética*, quien sugiere la existencia del libre albedrío. El segundo caso atiende a la propuesta de Libet basada en sus famosos experimentos neurocientíficos, en virtud de los cuales se plantea un determinismo conductual que implica la inexistencia del libre albedrío. El análisis de estos dos casos permitirá cumplir dos objetivos. El primero de ellos es ilustrar cómo se aborda la cuestión desde dos perspectivas cercanas, aunque diferentes, la psicológica y la neurocientífica. El segundo de ellos consiste en ver cómo puede llegar a afirmarse injustificadamente bien la existencia del libre albedrío o bien su inexistencia sin ofrecer una demostración empírica en torno al cumplimiento o incumplimiento de las condiciones de posibilidad del mismo.

Crítica a la perspectiva psicológica: libertad como libre-de y libre-para.

Muchos son los autores que han atendido a la cuestión de la libertad desde una perspectiva naturalista, incluso antes del auge neurocientífico. Antes de profundizar en los problemas que la neurociencia encuentra a la hora de abordar la cuestión del libre albedrío conviene atender a los problemas con los que se ha enfrentado la psicología. Para ilustrar la insuficiencia de esta rama del saber en torno al problema del libre albedrío atenderemos a Aranguren, concretamente al capítulo de su *Ética* titulado “La realidad constitutivamente moral del hombre: moral como estructura” (1994). Entre las estructuras psicobiológicas que este autor analiza, y que constituirían al ser humano como un ser estructuralmente moral, se encuentra la estructura de la libertad, a la que caracteriza en dos dimensiones. Atender a su propuesta permitirá comprender la dimensión psicológica más básica de la libertad, poniendo de relieve las dificultades que esta perspectiva encuentra para resolver el problema de si la conducta humana es libre o está determinada unívocamente. Tras ello profundizaremos en la cuestión dando paso a una crítica del mismo tipo, esta vez enfocada a las neurociencias, atendiendo al experimento de Libet y algunos de sus críticos.

Atenderemos aquí a una cuestión que ya hemos comentado en el epígrafe anterior, pero que es preciso retomar de nuevo. Según propone Aranguren, las estructuras

psicobiológicas de nuestra especie configuran la libertad en dos dimensiones, libertad como *libre-de* estímulos, en tanto que el ser humano escapa del esquema conductual estímulo-respuesta, y libertad como *libre-para*, en tanto que el esquema estímulo-cognición-respuesta abre, al menos aparentemente, un abanico de posibles respuestas ante un mismo estímulo.

Efectivamente, la libertad para preferir esta o la otra posibilidad procede de que la inteligencia como versión a la realidad concibe o proyecta diversas posibilidades de salida de la situación actual. (Aranguren, 1994: 51)

A lo cual añade que se trata “de una estructura a la vez compleja y unitaria”, siendo el caso que “la libertad no reposa sobre sí misma, sino que se levanta sobre las estructuras psicobiológicas” (Aranguren, 1994: 52).

Ante un estímulo, según indica el autor, el ser humano no tiene una respuesta predeterminada biológicamente, en ese sentido es *libre-de* los estímulos. La configuración de la respuesta precisa hacer uso de la inteligencia (Aranguren, 1994: 47). Nuestra especie presenta además unas predisposiciones biológicas que configuran unas tendencias o «ferencias» inconclusas, que permiten orientar la acción de forma probabilística, sin determinarla unívocamente.²⁷ En este punto es clave el papel de la cognición, a la que corresponde completar el proceso de respuesta al medio, conformándose así el esquema conductual estímulo-cognición-respuesta. Así, la elección de un curso de acción pone en ejercicio la otra dimensión de la libertad, la *libertad-para* preferir una acción concreta entre las posibles (Aranguren, 1994: 48).

Ahora bien, la propuesta de Aranguren en relación a esta cuestión no resuelve, sin embargo, el problema de si existe o no el libre albedrío. A pesar de que el ser humano en su conducta es *libre-de* estímulos y *libre-para* realizar diferentes acciones mediando con la cognición, podría suceder que los procesos cognitivos en interrelación con los elementos del medio que operan como estímulos configurasen una red de causalidad que determinase la conducta, estando la cognición sujeta a una cadena causal, y siendo la elección realizada, entre las diferentes posibilidades aparentes, la única respuesta posible

²⁷ En relación a esta cuestión cabe mencionar la idea de la moralidad estratificada en dos niveles, el de las estructuras biológicas, que son orientativas y compartidas por todos los miembros de la especie, y el de los contenidos, que corresponden al bagaje cultural adquirido a partir de las diferentes respuestas posibles ante un mismo estímulo. (Ayala y Cela Conde, 2001: 533).

en unas condiciones dadas. En tal caso la libertad existiría solo en apariencia, ser *libre-de* estímulos, tal como lo plantea Aranguren, significaría que los estímulos ambientales no serían los únicos que determinarían la respuesta unívoca, sino que además habría factores cognitivos que se sumarían a la determinación conductual, sin evitar así su univocidad, y sin escapar, por tanto, de la cadena causal. Ser *libre-para* significaría, a su vez, que existirían virtualmente varias posibles respuestas ante ciertos estímulos ambientales, pero no existiría la posibilidad real de dar una respuesta diferente a la que de hecho se da, pues los factores cognitivos sumados a los ambientales determinarían una respuesta unívoca y causalmente determinada.

Si bien hemos atendido a Aranguren por ofrecer un ejemplo paradigmático de la perspectiva psicológica en relación a la cuestión del libre albedrío, podemos afirmar que la psicología, en términos más generales, se presenta igual de infructuosa a la hora de desvelar la existencia o inexistencia de las condiciones de posibilidad de la libertad.

Merece la pena destacar que Aranguren acierta al señalar, aunque implícitamente, las mismas condiciones de posibilidad del libre albedrío que hemos señalado anteriormente, coincidiendo así con Cortina y Evers. Respecto a la condición *L1*, que refiere a la posibilidad de optar entre más de una opción a la hora de actuar, Aranguren atiende a ella al hablar de *libertad-de*, caracterizando al sujeto como *libre-de* univocidad, es decir, capaz de escoger entre más de una opción. En cuanto a *L2*, referente a la autodeterminación, Aranguren atiende a ella al hablar de *libertad-para*, caracterizando al sujeto como *libre-para* escoger por autodeterminación. Lamentablemente su propuesta no logra demostrar que estas condiciones de hecho se den, ni tampoco los presupuestos sobre los que se asientan. En cuanto a *L1* Aranguren da por sentado que al escapar del esquema estímulo-respuesta existe la posibilidad real de ofrecer diferentes respuestas ante las mismas causas, sin tener en cuenta que los factores cognitivos del esquema estímulo-cognición-respuesta pueden operar deterministamente haciendo que la conducta sea en todo caso unívoca y sujeta a una cadena causal determinista. En cuanto a *L2* Aranguren da por supuesto que en el proceso cognitivo que articula la respuesta opera un *yo* que permite que la decisión y la acción se den por autodeterminación, sin embargo no demuestra la existencia de ese *yo* ni de su capacidad de autodeterminación.

Crítica a la pretensión neurocientífica de cuantificar la libertad.

Tomando el relevo a la psicología algunos neurocientíficos se han propuesto resolver el problema desde su campo de investigación. Cabe preguntarse, no obstante, si tal cosa es posible. El estudio neurocientífico de la cuestión nace de manos de los experimentos de Konhuber y Deecke (1965; 1989), pero será Libet quien abra un debate polémico en torno al tema. Tratando de apoyar la tesis de que existe el libre albedrío sus experimentos neurocientíficos sugirieron lo contrario. Su conocido experimento trata de medir si la decisión consciente es anterior o posterior a la actividad cerebral que posibilita su ejecución. Se pretendía mostrar que la decisión consciente de realizar una determinada acción vendría dada antes que el proceso neuronal de poner en marcha esa acción. Sin embargo, el experimento identifica una carga eléctrica que aparece con anterioridad a la decisión consciente, precediéndola en unos milisegundos. Lo que el autor concluye es que esa carga, a la que llama *potencial de disposición*, impulsa y en cierta medida determina la decisión con anterioridad a que esta aparezca en la conciencia. Por tanto, el proceso de la voluntad se iniciaría inconscientemente, solo unos milisegundos después tendríamos conciencia de él. No cabría hablar de acciones libres en la medida en que el sujeto no las elige conscientemente, según plantea el autor. Pudo matizar más tarde el resultado de sus experimentos, sugiriendo que, a pesar de que la decisión parte de forma inconsciente desde el cerebro, la voluntad puede vetarla, existiendo así cierto terreno para la libertad y la responsabilidad en la acción. (Libet et al., 1983; Libet, 1999)

La propuesta de Libet ha generado todo tipo de reacciones, entre las cuales podemos diferenciar dos grandes grupos. Por un lado, existen quienes ven en su propuesta una confirmación de que la conciencia es una manifestación del cerebro, siendo una ilusión nuestra autopercepción como sujetos libres. Ligada a esta postura encontramos aquella que considera el cerebro como un mecanismo más de un universo determinista, sin escapar de la univocidad que esto supone, no habiendo por tanto espacio para la libertad. Por otro lado, sin embargo, se han producido una serie de reacciones que critican las conclusiones del experimento, considerándolas desproporcionadas en relación con lo que en realidad este puede demostrar empíricamente. Cabe atender a estas críticas para identificar algunos apuntes metodológicos relevantes y para poner en tela de juicio la posibilidad misma de resolver el problema desde una perspectiva científica y naturalista.

Lo que Libet propone es muy sugerente, pero las conclusiones del experimento, tal y como argumentan sus críticos, no son adecuadas partiendo de los datos empíricos obtenidos gracias al mismo. Su experimento y las numerosas críticas que ha recibido (Banks y Pockett, 2007; Cortina, 2012a: 164-172; Habermas, 2006: 162-164; Pallarés-Domínguez, 2016; Roskies, 2010) ponen sobre la mesa el problema, cada vez más acuciante, de aclarar cómo se ha de proceder a la hora de investigar la cuestión del libre albedrío desde el paradigma naturalista de las ciencias empíricas. Si la ciencia pretende abordar este tema hemos de preguntarnos cómo debe hacerlo y si está capacitada para resolver un problema de esta envergadura. Para esbozar una respuesta a esta pregunta señalaremos los puntos más relevantes que Cortina plantea, quien recoge a su vez las críticas que Habermas dirige contra las conclusiones del experimento. Estos puntos son los siguientes:

- El experimento plantea una situación artificial y unos presupuestos que hacen imposible su extensión a condiciones más generales de toma de decisiones.
- Nuestras decisiones siempre cuentan con una parte de planificación, no pueden medirse por instantes como Libet plantea. Y, por si fuera poco, los sujetos del experimento habían sido entrenados, pudiendo generar disposiciones mecánicas en el cerebro en relación al tipo de decisiones del experimento.
- Libet supone la existencia de un *yo*, en algún lugar del cerebro, que no se ha podido demostrar empíricamente (y que es uno de los presupuestos [L2a] de las condiciones de posibilidad del libre albedrío, tal y como hemos indicado en apartados anteriores).
- La decisión libre quizás no tenga que ser necesariamente consciente, como Libet propone, quizás baste con que sea virtualmente consciente. Podría actuarse de forma no consciente, queriendo actuar así de hecho, y después poder argumentar conscientemente el porqué de esa decisión.
- Le falta conexión interna con razones, las cuales son necesarias para que exista una deliberación que permita la libertad.
- Además, sostiene Cortina, es imposible decir desde la ciencia empírica que la libertad no exista, el método empírico solo puede tratar hechos que puedan experimentarse, afirmar que todo está determinado causalmente es una afirmación metafísica que escapa al alcance de cualquier ciencia.

Un análisis detallado de estos puntos permite configurar un listado de recomendaciones metodológicas que han de regir la investigación neurocientífica en torno al libre albedrío:

- Las condiciones de todo experimento son locales y artificiales, contando con unos presupuestos que imposibilitan la extensión de los resultados a contextos más amplios de toma de decisiones. Esto supone que los resultados de un experimento han de ser manejados con cautela, no extendiendo injustificadamente las conclusiones a contextos de decisión diferentes a los del experimento.
- Por el momento no existen criterios adecuados para la cuantificación de las decisiones. El proceso de decisión es amplio y complejo, de ahí que una de las críticas a Libet sea su pretensión de cuantificar la decisión en instantes. Encaja aquí también la crítica a Libet de no incluir las razones en la toma de decisiones en su experimento. El complejo proceso de la decisión se ve intervenido por las razones que llevan a tomar una y no otra decisión.
- El entrenamiento y la generación de disposiciones mecánicas son, entre otros, algunos factores que pueden adulterar el resultado de un experimento basado en la toma de decisiones. Hay que tener en cuenta estos factores y diseñar el experimento de forma pertinente.
- La investigación neurocientífica en torno a la libertad cuenta con algunos presupuestos que no han podido demostrarse empíricamente. Un ejemplo de ello es el supuesto de la existencia de un *yo*, que sería el decisor libre en caso de haber libertad.
- Es necesario aclarar conceptos básicos, como el de *decisión libre*. Mientras que Libet considera que la decisión libre ha de ser consciente, otros autores no lo consideran así.
- Que haya o no libertad es un hecho natural, pero no todo hecho natural es necesariamente cognoscible desde las ciencias empíricas. Encontramos aquí el problema de si las ciencias están capacitadas para ofrecer algún conocimiento sobre la existencia o no existencia de la libertad, o si, por el contrario, no son capaces de ello.

Libet no logra, y en un sentido más general tampoco la neurociencia, resolver si las condiciones de posibilidad del libre albedrío existen en el ser humano. Respecto a *LI*

(la posibilidad real de escoger entre dos cursos de acción diferentes) el experimento no ofrece datos suficientes para concluir que la conducta se produzca de forma unívoca en el marco de una cadena causal, sin embargo tal cosa parece darse por sentado al considerar que el potencial de disposición funciona deterministamente, ya que se produce antes de que aparezca la decisión en la conciencia. La posibilidad de que exista una contingencia causal o la posibilidad de que pueda decidirse de forma libre sin ser consciente son, entre otros, algunos escenarios que no deberían ser descartados. En cuanto a *L2* (la elección se ha de dar por autodeterminación) Libet tampoco logra demostrar que esta condición no se dé, pues no puede demostrar ni desmentir la existencia de un *yo*, que sería el decisor libre.

Esta infecundidad de la neurociencia en torno al problema del libre albedrío nos remite a reconsiderar el último de los puntos expuestos anteriormente. Este cobra especial relevancia ya que eleva la crítica no solo a cómo operar metodológicamente, sino a la suficiencia misma del método científico para resolver el problema del libre albedrío. Si, como Cortina argumenta, la ciencia empírica no está capacitada para ofrecer un conocimiento empírico sobre la existencia o inexistencia del libre albedrío, entonces las recomendaciones metodológicas expuestas anteriormente serían vacuas, pues el propio intento de conocer mediante las ciencias naturales si somos o no libres sería una tentativa condenada al fracaso.

Los problemas se plantean, en principio, por una cuestión de método. Quien desee recurrir a lo que ha venido a llamarse «método científico», propio de las ciencias naturales, tiene que investigar causalmente, [...]. Lo que no cabe en esa relación causal de acontecimientos observables, expresados en tercera persona, es o bien lo científicamente incognoscible o bien lo inexistente.

Reconocer que la libertad no puede ser conocida por este método es practicar algo así como un *agnosticismo de la libertad*, que es lo que profesó Kant en la *Crítica de la razón pura*: no podemos probar objetivamente la existencia de la libertad porque, si existe, no es el método empírico de las ciencias el que nos permite llegar a ella. Sin embargo, lo que no puede hacer el científico es afirmar que la libertad no existe porque no puede probar su existencia haciendo uso de su método. Una *afirmación semejante es metafísica, no científica*. El neurocientífico no puede permitirse un *ateísmo de la libertad*. (Cortina, 2012a: 187-188).

Más allá de la necesidad de establecer unos procedimientos metodológicos adecuados, Habermas, coincidiendo en esto con Cortina, considera la posibilidad de que la existencia o inexistencia del libre albedrío no sea algo medible por las ciencias

empíricas. Sobre esta cuestión nos advierte de que algunas cosas son, simplemente, incuantificables.

Como en todo *design*, también aquí surge la pregunta sobre qué es lo que se mide, además de la cuestión filosófica de entrada concerniente a qué cosas pueden ser sometidas a medición. (Habermas, 2006: 162)

Considerando que pueda haber una imposibilidad de las ciencias empíricas para resolver la pregunta de si somos o no libres, Cortina sugiere que abordemos la cuestión de forma más asequible, hablando de condiciones de la libertad, y no de causas. Lo que la autora propone, por tanto, es una sustitución de conceptos. Mientras que el concepto de *causa* es entendido en términos de necesidad, el concepto de *condición* tiene un carácter probabilístico. Esto permitiría evitar los problemas del método científico asociados al concepto de causa. La acción humana se ve siempre condicionada y limitada por una serie de factores, incluso en el caso de que existiese cierto espacio para la libertad. Hablar de condiciones, en lugar de causas, contempla estos factores, tomados como condicionamientos que limitan la conducta de forma probabilística, y evita a su vez los problemas derivados de hablar de la acción en términos de necesidad.

La cuestión de si somos o no libres sigue siendo hoy en día una pregunta sin respuesta. En primer lugar, la ciencia empírica sugiere que quizás sea posible resolver el problema a través de su método. Por ello, mientras los científicos mantengan su interés en torno a la cuestión, se ha de alentar una investigación que sea adecuada metodológicamente, solo de esta manera podemos esperar que cualquiera de los pasos que se den desde las ciencias empíricas se realicen en la dirección adecuada y sin excederse en sus competencias. Quizás una adecuada investigación científica pueda en el futuro arrojar algo de luz sobre esta cuestión tan problemática a día de hoy. Por otro lado, tal y como hemos expuesto, existe la posibilidad de que la ciencia no sea capaz de resolver el problema, simplemente porque no esté capacitada para ello. Por este motivo no es posible renunciar a la reflexión filosófica en torno al problema. Esta, cabe señalar, seguirá siendo necesaria incluso si la ciencia ofreciese respuestas al respecto. En el hipotético caso de que llegemos a conocer si de hecho el ser humano puede actuar o no con libertad, la filosofía moral habría de ocuparse de reflexionar acerca de cómo deberíamos gestionar ese conocimiento y cómo afectaría a nuestra comprensión del ser humano, de las

sociedades, de las instituciones, del derecho, de moral, de la ética y de conceptos como el de responsabilidad, deber y obligación moral.

Capítulo 5. Interdisciplinariedad y método en la neuroética.

En capítulos anteriores hemos atendido a algunas aportaciones de Patricia S. Churchland y de Kathinka Evers. Ambas plantean, desde diferente perspectiva, una propuesta que se reconoce como perteneciente específicamente a la disciplina de la neuroética, proponiendo una serie de recomendaciones metodológicas a las que es necesario atender. Sus propuestas presentan dos modos paradigmáticos, aunque bien distintos, de considerar las posibilidades y el método propios de la neuroética. El caso de Churchland resulta especialmente problemático desde una perspectiva crítica, ya que la autora hace una propuesta con un sesgo marcadamente naturalista, bajo el rótulo de *neurofilosofía práctica*. Evers, por su parte, trata de compatibilizar la labor neurocientífica con la filosófica, proponiendo una metodología interdisciplinar, bajo el rótulo de *materialismo ilustrado*. Atender a ambas será de gran utilidad para el desarrollo de una propuesta metodológica adecuada para el ejercicio de neuroética crítica en clave cordial.

5.1. Retos metodológicos de la neurofilosofía práctica

En *El cerebro moral* (2013), Patricia S. Churchland ofrece un detallado informe sobre los logros que la biología y la neurociencia han cosechado en lo que concierne a las conductas sociales y morales. La exposición que Churchland nos ofrece atiende mayoritariamente a datos científicos descriptivos sobre el surgimiento y evolución de las conductas sociales en diferentes especies de animales, prestando especial atención a los mamíferos. Atiende a la base cerebral que se asocia a tales conductas y a los cambios estructurales, funcionales y bioquímicos que acompañaron a esos cambios evolutivos. Por este motivo la mayoría de la obra de Churchland es un texto eminentemente científico. Sin embargo, la autora aborda cuestiones de calado metodológico y filosófico en algunos capítulos: en los capítulos 1 y 7 especialmente, así como en el capítulo 8, y en los últimos apartados de los capítulos 2 y 5. La reflexión filosófica interviene de forma cardinal en el trabajo de Churchland, de tal modo que la labor de las neurociencias es combinada con la reflexión de la filosofía moral. En este sentido su investigación se configura como una propuesta propia de la neuroética. Cabe preguntarnos si los argumentos que Churchland expone son correctos y someterlos a crítica allí donde veamos debilidades. Con

afirmaciones tan arriesgadas como que sí podemos derivar condicionales morales a partir de hechos, Churchland pone sobre la mesa el viejo tema de la falacia naturalista, esta vez puesta en cuestión. También abordará cuestiones relativas al establecimiento y seguimiento de normas.

El Capítulo 1 de *El cerebro moral* sirve como introducción a la obra y es de corte claramente filosófico. Al comienzo del capítulo Churchland cuenta que durante gran parte de su vida académica como filósofa evitó sumergirse en problemas de calado moral, puesto que durante años consideró estas cuestiones como desvinculadas de lo tangible, sin guardar estrecha relación con temas como la evolución y el cerebro. Afirma que esta ausencia de relación con lo tangible suponía el riesgo de flotar en un mar de simples opiniones, por convincentes que estas pudiesen parecer. No obstante, la neurociencia, la biología evolutiva, la psicología experimental y la genética ofrecen nuevos datos que plantean la posibilidad de vincular las bases de la conducta moral con esos elementos tangibles, respecto a los cuales las cuestiones morales parecían estar desconexas años atrás a ojos de Churchland. Ahora reconoce, desde estas disciplinas científicas, que la naturaleza social y moral del ser humano puede abordarse desde las ciencias empíricas. Churchland señala que las prácticas sociales y culturales, lo que compondría la esfera de la moral como contenido, no es el foco de interés de su obra, esta se centra más en la dimensión estructural de la moral, en su origen evolutivo y su funcionamiento neuronal. Esto no significa que la autora niegue la importancia de esa dimensión de la moral compuesta por contenidos históricos, contingentes y culturales. De hecho, aunque no de forma tan extensa, se ocupará de ello en algunas partes del libro.

Como puede apreciarse, el acercamiento de Churchland al estudio de la moralidad es naturalista. Será preciso analizar si este naturalismo queda restringido a las cuestiones puramente científicas, o si se adentra en problemas filosóficos como el de la fundamentación de lo moral, pudiendo resultar reduccionista. Respecto a ello Churchland habla del *cientismo* —sinónimo de *cientificismo*—, entendido como la pretensión de que la ciencia puede explicarlo todo. Asegura que las iniciativas de carácter científico no pretenden en absoluto desplazar a las artes o las humanidades. Sin embargo, en lo que concierne a la moral, dice, sí que existen campos en los que la ciencia debería desplazar a la filosofía más especulativa. Señala, por ejemplo, que las afirmaciones filosóficas acerca de la naturaleza de las cosas, como la intuición moral, son vulnerables. En los

casos en que la filosofía trata de explicar elementos cuantificables por las neurociencias, como el origen de la intuición moral y sus estructuras subyacentes, “la filosofía y la ciencia operan en el mismo campo, y los datos deberían superar cualquier reflexión especulativa” (Churchland, 2013: 14). Lo que Churchland sostiene es que la filosofía ha abordado históricamente cuestiones acerca de la naturaleza de algunas realidades de las que ahora la ciencia puede hacerse cargo, entre estas cuestiones encontramos el estudio de las condiciones naturales de la moralidad (Ayala y Cela Conde, 2001; Ayala, 2006; Joyce; Eisenberg, 2000; Tagny, 2007; Rizolatti, 2006; Morgado, 2010; Damasio, 2005; Damasio, 2008; Churchland, 2013; Goldberg, 2002; Forbes y Grafman, 2010; Libet 1999; Roskies, 2010; Prinz, 2001; Nichols, 2005; Waal, 1997, 2007; Brosnan y Waal, 2003; Tomasello y Vaish, 2012; Haidt, 2001; Greene; Casebeer, 2003; Mikhail, 2008). La ciencia ofrece datos empíricos en torno a esas realidades, mientras que la filosofía solo aporta reflexiones especulativas. Churchland declara abiertamente su interés por investigar los aspectos observables de lo moral, en tanto que hecho natural.

El naturalismo, además de alejarnos de las inferencias estúpidas, halla las raíces de la moralidad en cómo somos, en lo que nos ocupa y lo que nos preocupa, en nuestra naturaleza, en definitiva. Ni el suprarrealismo (los dioses sobrenaturales) ni un enrarecido concepto poco realista de la razón nos explica en qué consiste la placa base moral. (Churchland, 2013: 16)

La ciencia empírica opera en el paradigma naturalista, y en función de ello ofrece mejores soluciones que la mera especulación a la hora de explicar la naturaleza de ciertas realidades, en la medida en que tal naturaleza sea tangible mediante la experiencia. En estos casos el desplazamiento de la filosofía moral por parte de la ciencia está justificado, pues aquella pretendía abordar cuestiones que correspondían a las competencias de esta. Sin embargo, la filosofía moral aún conserva algunas competencias que le son propias. El problema en la propuesta de Churchland es que pretende establecer conexiones entre esa placa base moral, que concierne al nivel estructural y neurobiológico, siendo susceptible de explicación y descripción, y la normatividad moral, invadiendo así el ámbito justificativo respecto al cual la investigación científica está imposibilitada metodológicamente (Buchanan, 2009; Apel, 1985; Habermas, 2006; Aranguren, 1994; Gould, 1999; Hume 1988, 1953; Moore, 1983; Salles, 2013, 2016, 2018; Jácomo, 2011, 2013a, 2013b, 2014; Evers, 2010; Cartwright, 2010; Teehan y diCarlo, 2004; Cortina, 2012a; García-Marzá, 2016; Conill, 2012). Al no establecer una distinción clara entre la

base neurobiológica de la moral y los fundamentos normativos de principios y normas morales, Churchland corre el riesgo de cometer un reduccionismo naturalista.

Sobre la falacia naturalista.

Lo más problemático de la propuesta de Churchland es su afirmación de que en algunos casos, no en todos, es posible encontrar en la naturaleza elementos que resultan moralmente normativos. Volviendo la vista a Hume, Churchland pretende mostrar que existen casos en los que se pasa del *es* al *debe ser*, sin incurrir en reduccionismo naturalista. Por un lado, como naturalista, Hume quiso evitar la arbitrariedad de cierto tipo de deducciones, especialmente cuando el salto de los hechos al deber moral no estaba lógicamente justificado. Por otro lado, Hume halla las raíces de la moralidad en la naturaleza humana. ¿Cómo es posible, se pregunta Churchland, que la naturaleza establezca las bases de nuestra moralidad —en una dimensión explicable por la ciencia— pero no dé ningún tipo de orientación moralmente normativa sobre cuál debe ser nuestra conducta?

Veamos la crítica de Churchland en relación a Hume. Las intuiciones humeanas acerca de la naturaleza moral del ser humano eran acertadas en muchos de sus puntos, y supusieron un precedente en muchas cuestiones de las cuales las neurociencias y la biología pueden dar cuenta hoy en día. Efectivamente en la naturaleza humana, en su estructura biológica y neuronal, pueden identificarse las bases de la moralidad, la plataforma neuronal de la moral, tal y como Churchland refiere a ella. La dimensión biológica y neuronal de la moral, como dimensión estructural, es tangible, es objeto de las ciencias empíricas, y en algún sentido sí orienta la acción. El ser humano, así como muchos otros animales, siente una predisposición al cuidado de los cercanos, al altruismo dentro de su grupo, etc., y lo hace en función de su naturaleza biológica y neuronal (Ayala y Cela Conde, 2001; Ayala, 2006; Joyce; Eisenberg, 2000; Tagney, 2007; Greene; Damasio, 2005; Damasio, 2008; Haidt, 2001; Churchland, 2013; Waal, 1997, 2007; Brosnan y Waal, 2003; Tomasello y Vaish, 2012; Haidt, 2001; Greene). Sus impresiones de lo bueno y de lo malo, sus preferencias en torno a qué conductas sociales adoptar, etc., están en gran medida condicionadas —no determinadas— por esa naturaleza (Ayala y Cela Conde, 2001; Aranguren, 1994; Hauser, 2008). Ese sustrato natural condiciona su conducta y sus percepciones morales, orientando la acción en ciertas direcciones. Todo ello es materia de estudio para la ciencia empírica. Ahora bien, que la naturaleza

predisponga para ello, incluso que nos oriente en sentido prudencial, ¿significa que tales cursos de acción se ajustan a las conductas que *deben* adoptarse en sentido moral? Es decir, ¿en qué sentido esas predisposiciones pueden tomarse como prescripciones morales?, ¿qué las hace moralmente normativas? Es un exceso en sus competencias la pretensión de la ciencia de derivar un deber moral de esos códigos inscritos en nuestra naturaleza, aún si operan como guías de la acción.

No obstante, Churchland plantea que en cierto sentido sí que puede derivarse un *deber* de un *es*. Considera que dos han sido los principales hechos que han llevado a la falacia naturalista a adquirir entidad filosófica como fiable antagonista de un enfoque naturalista de la moral, sin merecer tal estatuto, según sostiene. El primero de esos hechos ha sido el de tomar la deducción del *ser* al *deber* como una derivación lógica, en la que un argumento es válido cuando la conclusión es indudablemente verdadera si las premisas de las que se sigue lo son, sin margen de maniobra. Churchland sostiene que el tipo de deducción del *ser* al *deber ser* puede ser una inferencia de carácter más flexible que el de una derivación lógica. El segundo hecho que señala es que una gran parte de la tradición filosófica, especialmente la kantiana, consideró que la biología tenía muy poco que decir respecto a la moral.

Hemos de analizar, por tanto, la sugerencia que Churchland hace en lo que concierne al estatuto lógico de la deducción naturalista de prescripciones morales. La autora propone que en sentido amplio es posible inferir lo que debe hacerse a partir del conocimiento, las percepciones, las emociones y el entendimiento. Pone varios ejemplos, entre ellos el siguiente: ante un espantoso dolor de muelas, puede deducirse que se debería ir al dentista. Frente a determinados hechos podemos extraer prescripciones conductuales en sentido prudencial. El hambre es un indicador de que debemos comer, el dolor invita a huir de aquello que nos lo causa, el sueño nos aconseja dormir, etc. Sin duda ciertos hechos nos sugieren qué tipo de acciones debemos acometer en un momento dado. Pero ese *deber* no es un *deber moral*. La confusión entre un deber prudencial y un deber moral es un error de falacia de ambigüedad por equívoco. En casos como este los términos son homónimos, pero no sinónimos (García Damboreana, 2000). Condicionales como “si queremos tener un sueldo debemos trabajar” y “si no queremos sentir dolor deberíamos alejarnos del fuego”, son ejemplos de deberes prudenciales, pero no morales. Estos ejemplos nos dicen qué acciones nos conducen a un fin, si pretendemos alcanzar ese fin

la prudencia nos aconseja qué acciones realizar, ya que son el medio adecuado para la consecución del fin que se persigue. Cabría preguntarse, en estos casos, si la naturaleza nos dice algo sobre la bondad moral de tales fines o sobre el carácter moral de tales prescripciones. La mera explicación naturalista que propone un fin como bueno no permite alcanzar una justificación de carácter fundamentador. Es cierto que algunos fines son considerados como algo bueno o malo en función de nuestra biología (Buchanan, 2009; MacIntyre, 2001; Nussbaum, 2006), que ha dispuesto tales percepciones como medios que faciliten nuestra supervivencia. Aquello que nos agrada suele hacerlo precisamente porque nuestra biología hace agradable lo que facilita nuestro éxito evolutivo, mientras que aquello que nos desagradaba lo entorpece. Ahora bien, que algo sea bueno para la supervivencia y la reproducción no significa que sea bueno o justificable desde el punto de vista moral.

La autora considera que la resolución de problemas en el mundo es uno de los ámbitos más relevantes en la dimensión moral, y en ese sentido la neurociencia y el estudio de la evolución biológica y cerebral tienen mucho que aportar. Sugiere adoptar una perspectiva científica y neurobiológica en torno a la resolución de problemas, la evaluación y la toma de decisiones en el entorno social, en el que además la biología preconfigura ciertos valores iniciales. En ese contexto los mecanismos neuronales que toman parte pueden identificarse y estudiarse por la ciencia empírica. La advertencia de Hume acerca del reduccionismo naturalista es desafortunada desde esta perspectiva científica, sostiene Churchland, sin presentar buenos argumentos para defender esta tesis. Mientras que los datos de la ciencia son descriptivos, aquellas prescripciones que esos datos puedan sugerir son más un consejo prudencial de carácter instrumental que la encarnación normativa de un deber en sentido moral.

El ser humano es un animal social e inteligente que ha de resolver problemas en relación con sus congéneres en el seno de su grupo. La resolución de estos problemas ha de atender necesariamente a los hechos para propiciar una conducta que resuelva los obstáculos que el individuo va encontrando en su entorno, tanto en el social como en otros (Colmenares, 2005). En ese sentido puede derivarse un condicional de un hecho, puede sostenerse que el individuo debe actuar de cierta manera. El procedimiento que permite inferir estas prescripciones a partir de los hechos no es de deducción, sostiene Churchland, estos problemas son problemas prácticos y sociales cuya resolución se basa en la

satisfacción de restricción. Esta implica factores diversos, de diferente importancia y probabilidades, que son barajados por el individuo hasta dar con una solución adecuada a la situación. Aunque esta no sea necesariamente la mejor forma de resolver el problema, es al menos una forma de resolverlo.

Churchland palpa la relación que aquí hay entre las estructuras y las conductas morales sujetas a normas, pero no hace un buen diagnóstico. Parte de esas estructuras de resolución de problemas componen la base neuronal y psicológica de la conducta moral. Esta se desarrolla como una respuesta al problema abierto de qué conducta debe adoptarse cuando existe la opción de escoger entre varias vías de acción posibles, en un entorno social en el que los resultados de esa conducta comprometen los intereses de otros individuos. Si bien la estructura de resolución de problemas es la misma, en gran parte, tanto para problemas morales como para casos prácticos moralmente neutros, los primeros tienen características propias y presentan su propia especificidad. La autorreflexión de la que es capaz el ser humano hace que este pueda plantearse cuáles son los motivos de la obligación y la norma en entornos sociales (Ayala y Cela Conde, 2001; Ayala, 2006), el ser humano es un animal inteligente y autorreflexivo cuyas facultades han abierto ante sí el problema de lo normativo (Koorsgaard, 2000). La dimensión moral del ser humano abre así el espacio de la ética, que puede diferenciarse de la resolución de problemas asociada a prescripciones prudenciales, en tanto que estos problemas pueden ser moralmente neutros. Churchland no logra distinguir que el tipo de condicional prudencial que prescribe un curso de acción como posible resolución de un problema práctico se distingue de aquellas otras prescripciones que indagan en el origen mismo de la normatividad y la obligación moral. En el primero de los casos será posible atender a los hechos para inferir cómo hemos de actuar, en el segundo no.

La propuesta de Churchland atiende a cómo nuestra biología marca nuestra conducta moral, en relación a ello hace un acertado diagnóstico al considerar que nuestra estructura moral ofrece valores y códigos que regulan nuestra conducta, reconociendo a su vez que estos pueden ser modificados en el proceso de endoculturación (Ayala y Cela Conde, 2001; Hauser, 2008; Evers, 2010; Aranguren, 1994; Piaget, 1987; Kohlberg, 1992; Cortina, 2012a, 2017), dando lugar a una serie de contenidos contingentes e históricos, aun si tienen a la base una serie de condicionantes biológicos comunes en todos los seres humanos. El error de Churchland reside en su análisis de la falacia naturalista.

Aunque ciertos hechos sugieran una serie de cursos de acción como adecuados cuando nos enfrentamos a esos hechos como problemas —ir al dentista si te duele una muela—, eso no significa que ese curso de acción se constituya como norma moral (“estás moralmente obligado a ir al dentista”), más bien se constituirá como consejo prudencial (“es aconsejable que vayas al dentista”), con un valor instrumental en aras de alcanzar algo que el sujeto considera como bueno (evitar el dolor). Ahora bien, que el sujeto considere o perciba algo como bueno, no significa que sea conocedor del bien como objeto natural (Moore 1983; Teehan y diCarlo 2004, 34-35; Hill 1976; Korsgaard, 2000). Una cosa es conocer los mecanismos biológicos y cerebrales que subyacen a nuestras percepciones de lo bueno y otra conocer el bien en sí.

A diferencia de otros neuroeticistas reduccionistas como Gazzaniga (2006) o Mora (2007), Churchland no está diciendo que los códigos cerebrales, biológicos y genéticos de conducta social deban ser tomados como normas morales. Lo que Churchland sostiene es que tales códigos sientan las bases estructurales de la conducta moral, suponiendo una conexión entre el *ser* y el *deber ser* en algunos casos concretos. Asimismo, reconoce el papel de las sociedades y las culturas en lo que hace a los contenidos morales. Churchland no cree que los contenidos morales sean totalmente contingentes e históricos, sino que componen variaciones y matizaciones de valores biológicamente condicionados. De este modo propone la existencia de unos valores primigenios o innatos, pero flexibles, que van maleándose en los diferentes contextos sociales, y que pueden encarnarse en diferentes normas.

La propuesta de Churchland en este punto presenta algunas virtudes, pero también defectos insalvables. La base estructural de la moral viene dada por nuestra naturaleza biológica, así que tiene sentido que algunas de nuestras percepciones sobre lo bueno y lo malo, algunos valores, nuestra percepción de lo justo, etc., vengán condicionados biológicamente (Ayala y Cela Conde, 2001; Damasio, 2005; Damasio, 2008; Churchland, 2013; Waal, 1997, 2007; Brosnan y Waal, 2003; Tomasello y Vaish, 2012; Aranguren, 1994; Hauser, 2008). En lo que refiere a nuestro propio bienestar y al de nuestros seres cercanos existe toda una serie de valores que vienen propiciados intuitivamente desde nuestra biología (Haidt, 2001; Greene). Si bien no se dan necesariamente de forma universal parece que sí que existe una propensión en forma de constante en las diversas culturas y momentos históricos. Ahora bien, que haya elementos biológicos que

condicionen nuestras intuiciones valorativas, o incluso códigos de conducta que desde nuestra biología orienten la acción, no significa que tales elementos hayan de ser adoptados como valores o normas morales (Cortina, 2012a, 2017). No es posible encontrar en la naturaleza objetos intrínsecamente normativos (Korsgaard, 2000), solo una aproximación filosófica y reflexiva a ellos es posible, desde ramas del conocimiento como la filosofía moral. Por ello los códigos de conducta que la evolución ha forjado en nuestra biología no pueden ser tomados como normas morales, sino únicamente como consejos prudenciales en aras de facilitar la supervivencia y la reproducción.

Sobre el problema de la falacia naturalista vuelve la autora en el último apartado del Capítulo 7 de *El cerebro moral*. Churchland dedica los capítulos precedentes —del 2 al 6— a exponer principalmente datos descriptivos acerca de la base biológica, neuronal, genética y evolutiva de las conductas sociales y morales en mamíferos, salvo dos apartados de crítica filosófica a Hauser y Haidt en el Capítulo 5 y una breve incursión filosófica en el capítulo 2. En el capítulo 7, de forma más extensa, se adentra en cuestiones eminentemente filosóficas. Lo que la autora se propone aquí es abordar la cuestión de las normas, las cuales considera estrechamente vinculadas con predisposiciones biológicas, aun si pueden presentar variaciones contingentes e históricas en función de la cultura o sociedad en la cual toman forma.

Churchland somete a crítica algunas de las perspectivas filosóficas que intentan dar solución a la cuestión de qué normas morales han de regir nuestra conducta. La crítica se enfoca en primer lugar a la Regla de Oro, tanto en su versión positiva —*trata a los demás como quieres que te traten a ti mismo*— como en su versión negativa —*no hagas a los demás lo que no quisieras que te hagan a ti mismo*—. Tras ello dedica dos apartados para someter a crítica y mostrar las debilidades de la propuesta normativa kantiana, en primer lugar, y la utilitarista y consecuencialista, en segundo lugar. En los dos siguientes apartados trata cuestiones introductorias para el que será el último apartado del capítulo, donde su propuesta toma forma, articulando una crítica a la falacia naturalista. Volveremos en el siguiente apartado sobre la cuestión de las normas, abordada de forma amplia, por el momento seguiremos analizando la cuestión de la falacia naturalista.

La autora señala que los códigos sociales de conducta, así como los datos científicos acerca de las bases neurobiológicas de la conducta moral, han sido tomados como datos relevantes a nivel descriptivo y antropológico, pero no han sido tomados

como centro de la moralidad en un sentido profundo y normativo. La distinción entre lo que *es* y lo que *debe ser* ha cristalizado en el campo filosófico como un proyecto normativo autónomo en relación a las descripciones de hechos. Churchland está en desacuerdo con esa pretendida autonomía de tal proyecto normativo, declara que “dadas sus convicciones sobre la autonomía del dominio normativo, muchos filósofos morales consideran la distinción entre hechos y valores como algo que echa por tierra toda la empresa de este libro” (Churchland, 2013: 204), lo cual denota que la autora se declara a favor de una conexión de inferencia entre hechos y deberes morales. Tal convicción está en un error, ya que la crítica filosófica sobre el reduccionismo naturalista no niega el valor de las aportaciones científicas en lo relativo a la conducta moral y sus bases estructurales (Moore, 1983; Teehan y diCarlo 2004).

A lo largo de este apartado Churchland somete a crítica el argumento que Moore sostuvo para defender que la deducción de un *deber moral* a partir de un hecho implica incurrir en una falacia. En términos generales lo que Moore sugiere es que identificar propiedades naturales o conjuntos de estas con lo bueno, malo o valioso es falaz (Moore, 1983: 9). Sostiene Moore que no hay respuesta a la pregunta de qué propiedades naturales se corresponden con lo bueno, lo correcto, lo valioso o lo justo. Estos elementos no son naturales y solo pueden abordarse especulativa e hipotéticamente por los filósofos, no empíricamente por los científicos.

Ahora bien, Churchland pone en entredicho esta tesis, ¿realmente no puede dirimirse qué es bueno, valioso o correcto moralmente atendiendo a la naturaleza de los seres que realizan y sufren las acciones? Si bien Churchland no lo expone explícitamente, su propuesta se compone de una doble vertiente filosófica, utilitarista por un lado y prudencial en sentido aristotélico por otro. Lo que está implícito en su propuesta es que al fin y al cabo lo bueno y lo malo no es algo intangible, sino que es bueno o malo en función de cómo afecta a aquellos que sufren las consecuencias de las acciones, bien sean sus propias acciones o las de otros. Así la felicidad, el bienestar, la salud y la supervivencia componen en términos generales una percepción de lo bueno compartida por la inmensa mayoría de seres humanos, así como por algunos otros animales (Buchanan, 2009; MacIntyre, 2001; Nussbaum, 2006). Asegura que “nuestras percepciones están impregnadas de valor, y en este sentido actúan como valencias” (Churchland, 2013: 207).

Esta perspectiva entronca en cierto sentido con el enfoque de las capacidades y con la idea de florecimiento que de manos de Amartya Sen (1979), MacIntyre (2001) y Nussbaum (2006) toma forma. Según esta perspectiva se han de observar por un lado las capacidades de los diferentes individuos en determinados contextos sociales y ecológicos, y por otro se ha de atender a su biología, que determina qué constituye un adecuado florecimiento para ese individuo concreto, en función de la especie a la que pertenezca e incluso de la cultura de la que forma parte. Desde esta perspectiva la satisfacción de determinadas cotas de bienestar, salud, desarrollo, etc., componen un bien en sí mismo, por naturaleza. Lo que Churchland defiende es que ese tipo de bienes naturales sirven para hacer prescripciones morales en algunas ocasiones. Si bien la autora reconoce la existencia de casos de falacia naturalista, no todos los casos en que se infiera un *deber moral* de un hecho incurren en esta falacia, asegura Churchland.

Lo valioso de su propuesta es que lleva a cabo un acercamiento entre hechos y normas. Efectivamente existen una serie de bienes naturales, cuya consecución ha de orientar en sentido prudencial la conducta de los individuos implicados. El adecuado desarrollo de un individuo, su florecimiento, su bienestar, incluso su felicidad, dependen de la consecución de esos bienes en gran medida. La filosofía moral tiene en cuenta este hecho (Moore, 1983; Teehan y diCarlo 2004), todas las teorías éticas observan de uno u otro modo que la aplicación de la norma garantice unos mínimos aceptables en lo que concierne a la consecución de esos bienes naturales. La universalidad del imperativo categórico (Kant, 1992, 2000), la búsqueda de la máxima utilidad (Bentham; Mill; Singer), la validez de una norma que se da por consenso en un contexto dialógico (Habermas; Apel; Cortina), por poner algunos ejemplos, atienden en última instancia a garantizar unas cotas mínimas de bienestar, felicidad y florecimiento de los afectados.

La ciencia puede hacer grandes aportaciones en lo que concierne al conocimiento de esos bienes naturales y a las prescripciones prudenciales que facilitan su consecución. Esos bienes son cuantificables, en tanto que percibidos a nivel biológico y neuronal, así como en relación a su valor biológico en la medida en que sean favorables para el organismo. La ciencia puede ayudarnos a comprender qué debemos hacer para tener una mejor salud o para mantener unas cotas de bienestar adecuadas, pero ese es un *debemos* en sentido prudencial, como resolución de un problema práctico, bajo criterios de satisfacción de restricción, no es un *deber* en sentido moral. La confusión entre un tipo

de deber y el otro hace a Churchland incurrir en una falacia de ambigüedad por equívoco, conducente a un error lógico en la argumentación por transposición categoría. Volveremos sobre esta cuestión en el Capítulo 6. La filosofía moral aborda cuestiones de un calado mucho más profundo, intangibles para la ciencia. Qué es el bien, qué es la justicia, qué papel juega la responsabilidad en la acción moral, qué normas han de regir la conducta para que ésta sea ética, de donde emana lo normativo y la obligación moral, etc. Todas ellas son cuestiones que exceden en mucho las competencias de la ciencia. Por ello, si bien cabe reconocer que Churchland realiza un acercamiento entre hechos y cierto tipo de prescripciones prudenciales, no logra argumentar por qué esos hechos implican obligaciones para el agente moral. La ciencia puede dar cuenta explicativa y descriptiva de la moral, pero no justificativa en sentido normativo (Buchanan, 2009; Apel, 1985; Habermas, 2006; Aranguren, 1994; Gould, 1999; Hume 1988, 1953; Moore, 1983; Salles, 2013, 2016, 2018; Jácomo, 2011, 2013a, 2013b, 2014; Evers, 2010; Cartwright, 2010; Teehan y diCarlo, 2004; Cortina, 2012a; García-Marzá, 2016).

Sobre las normas.

Además de abordar la cuestión de la falacia naturalista, en el Capítulo 7, que se titula «No como norma», Churchland reflexiona en torno a la cuestión de las normas. Esta cuestión no podía ser de mayor interés para nuestra investigación. El motivo de que la autora no se ocupe de la cuestión de las normas hasta el penúltimo capítulo de su libro se debe a la propia estructura de su proyecto. La resolución de problemas sociales puede cristalizar, si bien no siempre ocurre, en la constitución de reglas específicas. Éstas se producen como resultado de unos valores compartidos, que tienen origen biológico y cultural y que anteceden a las normas en la medida en que la biología predispone a ellos y la cultura los modifica y perpetúa, muchas veces de forma implícita. Las normas pueden tomar la forma de leyes, reglas, mandatos, etc., pero también pueden manifestarse de forma implícita en las prácticas sociales, en ambos casos pueden producir como resultado la generación de hábitos y guiar la acción (Waal, 1997, 2007; Brosnan y Waal, 2003; Tomasello y Vaish, 2012; Kohlberg, 1992; Piaget, 1987). Todas ellas se producen como resultado de problemas que surgen en la relación de unos individuos con otros en el seno de grupos y sociedades en los que se trata de mantener una convivencia estable.

El objetivo de Churchland en este apartado es someter a crítica las tentativas de proponer normas morales de conducta universalmente válidas, capaces de guiar nuestra

praxis. Pretende además explicar el modo en que los seres humanos son capaces de valorar la bondad y la justicia de ciertas normas y del seguimiento de las mismas. En su análisis de la cuestión, además, denunciará algunos problemas como el conflicto entre normas o la necesidad de imponer excepciones a las normas.

La crítica adopta tres focos principales: la regla de oro, el utilitarismo y el consecuencialismo, y el imperativo categórico kantiano. Churchland señala que desde estas tres perspectivas se busca el establecimiento de normas universales y sin excepciones, este será el aspecto de ellas respecto al que disiente.

En relación a la regla de oro esta encarna una de esas normas que pretende ser defendida sin excepciones y teniendo validez universalmente. En su versión positiva la regla enuncia que se ha de hacer a los demás lo que nos gustaría que nos hiciesen a nosotros mismos —si nos encontrásemos en la situación del otro—, en su versión negativa la regla prescribe no hacer a los demás aquello que no deseáramos para nosotros. Lo cierto es que la regla de oro sirve como guía para resolver ciertos problemas morales, pero presenta también unas limitaciones que han de ser tenidas en cuenta. Churchland se remite a algunos ejemplos de aplicación donde se hace patente la insuficiencia de la norma. Por ejemplo, nos expone cómo un jurado podría desear mandar a un acusado a prisión, sin que los miembros del jurado desearan eso para ellos mismos aun siendo culpables del mismo delito. También en el mundo de los negocios y el comercio uno podría desear para los demás lo que no desea para sí mismo. Asimismo, en los casos en que las necesidades de mi propia familia entran en conflicto con la ayuda prestada a los demás, se deseará para sí y para los suyos un trato preferencial, que denota también una insuficiencia en la aplicación de la regla de oro. También en casos de donaciones de órganos la aplicación de la norma es insuficiente, podría desear que alguien me donase un riñón si lo necesito, y eso no implica que haya de donar yo uno de los míos a quien lo necesite. Finalmente, Churchland expone una serie de casos aún más problemáticos, en los que alguien puede querer algo para sí y para otros, sin que esos otros en absoluto lo deseen, por ejemplo ser convertido a una secta, no lo deseo en absoluto para mí ni para otros, pero muchos devotos bienintencionados sí lo desean para sí y para los demás. Ello pone de manifiesto que cuando las dos partes presentan valores y concepciones de la realidad divergentes la aplicación de la regla de oro se vuelve muy problemática, en tanto que se relativiza lo que cada uno puede desear para sí y para otros.

También esgrime Churchland una crítica contra el utilitarismo y el consecuencialismo, cuestionando la maximización de la utilidad como principio moral rector de la acción. Las virtudes de la propuesta utilitarista son varias, y han de ser reconocidas a pesar de que tiene elementos que suponen un foco posible de crítica. El utilitarismo busca la consecución de la máxima felicidad para el mayor número, aplicándose el principio de utilidad de forma imparcial. Las diferentes formulaciones del principio de utilidad y su aplicación han podido evitar algunas de las críticas asociadas al daño que las minorías sufrirían en aras de propiciar la felicidad del mayor número. Por ejemplo, estableciendo la prioridad y la mayor valencia de ciertos intereses de un individuo frente a los intereses que suponen un daño marginal para una mayoría, por ejemplo dar una pequeña cantidad de dinero a una persona pobre de forma que cubra sus necesidades básicas sería más recomendable que repartir ese dinero entre una mayoría, para la cual el aumento de felicidad resultante es marginal.

Lo verdaderamente valioso, indica Churchland, es que el utilitarismo tiene en cuenta las consecuencias a la hora de valorar un plan, una norma o una acción. Tener en cuenta qué consecuencias sociales de felicidad y bienestar se derivan de las acciones que barajamos realizar es sin duda algo necesario a la hora de establecer normas morales. Sin embargo, es un error considerar que solo el utilitarismo hace esto, o que lo hace de forma más eficiente que otras propuestas éticas. La ética del discurso, de cuya tradición es heredera nuestra propuesta de una neuroética de la razón cordial, también tiene en cuenta las consecuencias, en la medida en que lo que busca es aplicar una norma que tenga en cuenta las razones posibles que defenderían los afectados de dicha norma, de forma que estos pudieran asentir ante ella, en el caso de que las consecuencias que se derivasen de la norma respondan a sus intereses, minimizando el conflicto con los intereses ajenos en la medida en que el establecimiento de la norma se produciría por consenso. Si cada individuo puede exponer en una comunidad de comunicación, mediante el ejercicio del diálogo, qué normas y acciones promueven su felicidad, aportando razones para ello, el utilitarismo debería reconocer en la ética dialógica uno de los medios más eficientes para la maximización de la felicidad del mayor número, que mejoraría las cotas de felicidad en un cálculo de utilidad. El proceso dialógico de establecimiento de normas cumpliría con el principio de utilidad en la medida en que ayuda a promover la felicidad del mayor número, estando guiado por un criterio de universalidad.

El problema del utilitarismo es que el principio de utilidad se excede en sus pretensiones de aplicación. Aplicar este principio de forma rigurosa obligaría a llevar a cabo una serie de acciones que difícilmente podríamos aceptar. Llevar a los propios hijos a un colegio público, y no privado, para financiar la educación de otros niños, no ir de vacaciones para cubrir gastos básicos de personas que lo necesitan, ir privándonos de pequeños lujos para solventar los problemas de aquellos que están más desfavorecidos, etc., son desde luego acciones loables, pero es problemático argumentar cómo podríamos estar obligados moralmente a limar todo desnivel de bienestar propio para favorecer la mejora del bienestar ajeno. Si bien el principio de utilidad puede servir como guía de la acción, en el sentido en que lo hacía la regla de oro, su aplicación rigurosa plantea problemas que la propia teoría no puede resolver sin renunciar al principio de utilidad como principio rector de la acción, como norma de aplicación general que no admite excepciones.

La crítica en relación a la cuestión de las normas también es dirigida por Churchland hacia el imperativo categórico kantiano. Este es propuesto como un procedimiento para establecer normas completamente incondicionales y sin excepción de forma imparcial. La propuesta kantiana plantea que la facultad de la razón, en su uso práctico, es el principio rector de una buena voluntad, capacitando así al sujeto trascendental para ser autónomo moralmente, dándose a sí mismo normas morales con validez universal (Kant, 1992, 2000). La crítica de Churchland se articula en torno a dos puntos, el primero de ellos trata de poner sobre la mesa ejemplos en los que la aplicación del imperativo categórico no da como resultado lo que Kant promete, el segundo denuncia que el desapego de las emociones en el ejercicio de la razón pura en su uso práctico es, por diferentes motivos, problemático. El análisis de esta cuestión resulta especialmente relevante para nuestra investigación en la medida en que una neuroética de la razón cordial se enmarca en las éticas procedimentales herederas del pensamiento kantiano.

En lo que concierne al primer punto, la razón pura aplicada de forma coherente y estricta puede dar lugar a que defensores de ciertos fanatismos y fascismos consideren como susceptible de aplicación universal una máxima que en absoluto puede reconocerse como norma moralmente correcta.

Esto es en realidad el simple reconocimiento de que muchas personas creen que existen cosas peores que la muerte —el deshonor, ir al infierno, sufrir un dolor mortificante, etcétera—. [...] el

filtro kantiano no tiene más remedio que dejar pasar a muchos fascistas coherentes con sus creencias así como a multitud de fanáticos de la moral racionalmente coherentes consigo mismos.

Estos problemas indican que tomar la racionalidad pura y la coherencia como entramado de la moralidad es un camino equivocado. (Churchland 2013: 192-193)

Las dificultades que este problema plantea se ven resueltas por los procedimentalismos discursivos y dialógicos que, entroncados en la tradición kantiana, ponen en la esfera pública el dominio de la reflexión en torno al establecimiento de normas. En la medida en que las razones son aportadas y debatidas en condiciones que aspiran a un ideal contrafáctico regulativo, la situación ideal de habla, se evitará el problema de considerar como universalizable lo que uno cree adecuado bajo reflexión privada. La publicidad de la reflexión garantizará que el asentimiento ante una norma se produzca en términos de universalidad dado que el diálogo mismo tiene un carácter de inclusión universal en relación a los afectados por la norma. Los problemas específicos de una autonomía moral monológica quedan así resueltos trasladando esa autonomía moral al plano dialógico e intersubjetivo (Apel, 1985; Habermas, 2006, 2010; Cortina, 2007, 2009). Lamentablemente Churchland no alude a ello en este apartado. No obstante, en otros puntos de la obra encontramos momentos en que la autora reconoce el valor de la deliberación pública a la hora de establecer normas. En el primer capítulo podemos percibir esta cuestión atendiendo a afirmaciones como la siguiente:

[...] hallar soluciones óptimas a los problemas sociales suele requerir una gran dosis de sabiduría, buena voluntad, capacidad de negociación, conocimiento histórico e inteligencia. (Churchland, 2013: 16)

La buena voluntad, el conocimiento histórico y la capacidad de negociación son tres elementos que denotan la necesidad de tratar los problemas sociales en el ámbito público en aras de ofrecer soluciones óptimas para todos los afectados. Dado que los problemas morales son un tipo de problema social, la cita anterior es aplicable también a la resolución de estos. La reflexión y el asentimiento ante normas será insuficiente si se ejerce de forma privada, la única garantía de orientar universalmente la aplicación de las normas ante las que asentimos es que se procure garantizar la participación de todos los afectados en los procedimientos de establecimiento de normas, argumentando dialógicamente y no monológicamente.

La autora vuelve a comienzos del Capítulo 7 sobre esta cuestión, esta vez hace referencia a las habilidades precisas para llevar a cabo negociaciones de carácter público, queda expresado solo implícitamente que este tipo de negociaciones son un procedimiento adecuado para establecer normas:

Requiere tener habilidades complejas, incluida la habilidad para tratar de un modo efectivo con el desorden y la inestabilidad social, anticiparse a las consecuencias de un plan y predecir nuevos problemas, así como la habilidad para negociar productivamente sobre las normas explícitas y las instituciones. (Churchland, 2013: 182)

Por otra parte, la segunda crítica que Churchland hace al imperativo categórico kantiano se centra en los problemas que se derivan del ejercicio de una razón pura despojada de todo factor emocional. Churchland sostiene que “la razón no crea valores, sino que se configura en torno a ellos y los lleva hacia nuevas direcciones” (Churchland, 2013: 184). Coincidimos con Churchland en que el ejercicio de la razón pura, despojada de todo factor emocional, puede conducir a casos, como los citados anteriormente, en que un razonamiento coherente puede proponer normas moralmente problemáticas. Esta es precisamente la crítica que una ética de la razón cordial vierte sobre las filosofías procedimentales y kantianas de corte excesivamente racionalista, la razón opera siempre de forma encarnada, desde individuos concretos, en contextos sociales y comunicativos concretos, por lo cual el reconocimiento y el estudio de los factores emocionales implicados en los procedimientos de generación de normas no pueden ser desatendidos (Cortina, 2009). Coincidimos con Churchland en la necesidad de atender a la dimensión emocional en los procesos de establecimiento de normas morales, pero disentimos con ella respecto a la idea de que basta un estudio naturalista de estos procesos para descubrir criterios que permitan fundamentar las normas resultantes de esos procesos.

Además, expone, la evolución nos ha dotado de las emociones para poder procesar nuestra relación con el entorno y dar respuestas adecuadas a este. Las emociones han existido como elemento cognitivo de mediación entre el entorno y la respuesta conductual a este durante cientos de millones de años. La razón tal y como se da en los seres humanos, con un carácter autorreflexivo y abstracto, hizo su aparición hace tan solo unos cientos de miles de años. La razón juega, en sentido biológico, un papel complementario al de las emociones, enriquece, corrige, orienta, potencia e inhibe, si es necesario, los elementos emocionales (Darwin, 1979; Ayala y Cela Conde, 2001; Damasio, 2005; Damasio, 2008;

Churchland, 2013; Waal, 1997, 2007; Brosnan y Waal, 2003; Tomasello y Vaish, 2012; Aranguren, 1994; Hauser, 2008; Haidt, 2001; Greene; Cortina, 2009, 2012a, 2017). El ejercicio de la razón desvinculada de todo factor emocional estará dejando de lado aspectos importantes para la reflexión. A la hora de resolver problemas, la razón ofrece soluciones mejores si opera conjuntamente con los factores emocionales relevantes para el caso. Tal y como propone una ética de la razón cordial (Cortina, 2009) en la medida en que los afectados de una norma quieran participar en la deliberación, deberán de tener en cuenta las emociones y las capacidades estimativas que se ven implicadas en aquello que consideran valioso y deseable, un ejercicio puramente racional del diálogo desatiende elementos que juegan un papel crucial en la discusión.

Algunas investigaciones en neuroética y neuropolítica apoyan esta tesis, tales como las llevadas a cabo por Gorge E. Marcus (2002) y Sharon R. Krause (2008) en *The sentimental citizen: Emotion in democratic politics* y *Civil passions: Moral sentiment and democratic deliberation*, respectivamente. Ambas obras encarnan un amplio proyecto de crítica a las políticas asentadas sobre tradiciones de deliberación pública de corte racionalista.

Krause dirigirá la crítica en el primer capítulo de su obra contra la propuesta habermasiana y la rawlsiana, por ser excesivamente racionalistas. La autora reivindica el papel de las emociones en la política democrática, para ello presenta dos puntos clave, una especial atención a las pasiones y un revisionismo en la política. Ese revisionismo se vierte sobre las éticas dialógicas y discursivas, lo cual hace especialmente interesante su propuesta para nuestra investigación, en la medida en que guarda una notable afinidad con la propuesta de un *ethica cordis*. Krause busca una reivindicación de una justicia afectiva pero imparcial, para ello es precisa una revisión de las políticas democráticas y de los derechos liberales, especialmente en lo que concierne a la razón pública. Critica a Habermas y a Rawls como ejemplos paradigmáticos de una filosofía política asentada en una posición racionalista, especialmente en lo que concierne a la cuestión de la justicia. Les acusa de presentar una carencia en lo que corresponde a las emociones, la cual impide abordar adecuadamente la cuestión de la imparcialidad y de la justicia. En el capítulo tercero lleva a cabo un contraste entre la noción de deliberación humeana y la kantiana, destacando que en aquella el papel de las emociones y el de la razón no pueden darse de forma aislada, sino que operan conjuntamente. En los siguientes capítulos profundiza en

esa comprensión de los otros. Sostiene que la reciprocidad y la deliberación pública solo pueden ser posibles en la medida en que se tenga en cuenta a los otros, no solo en términos racionales, sino también teniendo en cuenta aspectos emocionales y afectivos. En la medida en que las emociones sean relevantes para la reciprocidad y la deliberación pública la razón por sí sola será insuficiente, la inclusión de lo emocional en la política será necesaria, y solo a través de ella se podrá alcanzar una verdadera imparcialidad.

La deliberación en sociedades democráticas no podría hacerse de forma desconexa a las emociones, así que lo más conveniente no sería trascender más allá de las pasiones, sino reformar la política de forma que esta esté habilitada para una deliberación racional que pueda ser partícipe de las pasiones. Ensanchar nuestra comprensión de los otros y buscar una fusión de horizontes es algo necesario para la deliberación pública, y solo puede hacerse atendiendo al papel de las pasiones, a la comprensión tanto de las propias como de las ajenas. Las críticas de Churchland y Marcus encuentran acogida en la propuesta de Krause, que denuncia y ofrece soluciones para la deficiencia que presentan algunas propuestas de la ética dialógica, concretamente aquellas que se articulan bajo parámetros excesivamente racionalistas.

Si bien la razón de este tipo de éticas, así como de las políticas deliberativas que son su correlato político, es una razón descentralizada, se precisa la inclusión de lo emocional y su reconocimiento, de modo que sea tenido en cuenta de forma explícita a la hora de deliberar. Ese es precisamente el núcleo de una ética de la razón cordial (Cortina, 2009), cuya propuesta reconfiguraremos para proponer una neuroética crítica que compatibilice cordialmente la razón y la dimensión emocional.

A pesar de esta crítica, que reivindica el papel de las emociones, no hemos de caer en el error de subestimar el papel y la importancia de la razón. Churchland, en la última parte del Capítulo 6, en el que investiga detenidamente algunos aspectos de las bases genéticas y cerebrales de la conducta, pone de relieve el carácter no reflexivo que en ocasiones presenta el juicio moral. Para ello atiende a los psicólogos Hauser y Haidt, de quienes hemos hablado en detalle en capítulos anteriores. El primero de ellos expone cómo nuestra conducta moral parte de unas estructuras básicas que orientan el desarrollo de las facultades morales y la adopción de contenidos, como una gramática moral universal en el sentido que Chomsky expuso en relación al lenguaje (1987). Estas estructuras morales predisponen, entre otras cosas, para hacer juicios morales de forma

no consciente, aunque esto no sea aplicable a todos los casos en que se articula un juicio de este tipo (Hauser, 2008). En cuanto a Haidt, este defiende la existencia de un intuicionismo social, argumentando que la intuición influye en la resolución de problemas sociales, en la conducta moral y en el juicio moral (Haidt, 2001, 2013; Greene y Haidt, 2002). Estas intuiciones vendrían asociadas biológicamente a situaciones y estímulos en que resultaban adaptativas, aunque culturalmente pueden configurarse y desarrollarse de forma variable. Ambas propuestas ponen de relieve que existe un estrato no reflexivo y que no se muestra explícitamente ante la conciencia en el juicio moral, la toma de decisiones y la conducta moral. El papel de la razón será necesario para revisar esas intuiciones morales y poder dirimir si existen o no razones que justifiquen el juicio basado en la intuición.

5.2. Materialismo ilustrado como método en neuroética

Kathinka Evers ha realizado interesantes investigaciones en el campo de la neuroética. Su propuesta en *Neuroética, Cuando la materia se despierta* (2010) es especialmente interesante porque adopta una actitud crítica en relación a la psicología y la neurobiología precedente, valiéndose para ello de las más recientes investigaciones neurocientíficas, siendo especialmente deudora de Changeux (1985, 1993, 2001, 2010), a quien cita y refiere repetidamente a lo largo de la obra. Su concepción del cerebro como un órgano plástico, emotivo, variable y creativo permite a la autora alejar las antiguas concepciones mecanicistas para proponer un modelo del cerebro como una materia dinámica y despierta, en el más literal sentido de la palabra. Se adentra en nociones fundamentales para la investigación de lo moral, como son la de la conciencia, la de la identidad, la del yo, la de la responsabilidad personal y la de la libertad, así como cuestiones relativas a la base biológica y neuronal de nuestra conducta moral y cómo la cultura influye en ella. Evers llevará a cabo una propuesta de investigación científica que exige programas cargados de responsabilidad y que salvaguarden los objetivos de la ética. Así, sometiendo a reflexión filosófica el saber que las neurociencias nos ofrecen, esta autora hace un ejercicio ejemplar de lo que podemos llamar *neuroética*. Atender a ella nos servirá como un caso paradigmático del ejercicio de la neuroética, en el que nos ofrece pautas metodológicas muy interesantes, permitiéndonos además entrar en las cuestiones principales que se abordan en esta disciplina. En el presente epígrafe no atenderemos al planteamiento que Evers hace en torno a la cuestión del libre albedrío, en el Capítulo 2

de su obra, ya que hemos atendido de forma detallada a ello en el segundo epígrafe del capítulo anterior de nuestra investigación.

A lo largo del primer capítulo plantea y caracteriza el método que ella considera más adecuado para proceder en la investigación en neuroética. Llama a este método *materialismo ilustrado*. Asimismo, sitúa la conciencia y la cognición —incluiremos en esta las emociones²⁸— como objetos de estudio adecuados para las ciencias empíricas, y en especial para las neurociencias. Este hecho es especialmente relevante para Evers, ya que históricamente, explica, toda referencia a términos mentales ha sido considerada como “no científica” o acusada de psicologismo (Frege, 1892). Además, gracias a la posibilidad de estudiar las conductas morales, la conciencia y la cognición a través del cerebro las neurociencias adquieren una *pertinencia normativa*, asegura. Esta pertinencia normativa no es reduccionista en los términos en que Evers la describe, tendremos ocasión de atender a ello en detalle.

Uno de los primeros interrogantes que Evers plantea en su obra es cómo las ciencias naturales pueden profundizar nuestra comprensión del pensamiento moral mediante el estudio del cerebro. El planteamiento de esta cuestión responde en realidad a la búsqueda del método adecuado mediante el que las neurociencias han de proceder en la neuroética. A este respecto cabe plantear las siguientes cuestiones: ¿qué método es el adecuado para el estudio neurocientífico de lo moral? ¿Qué objetos pueden ser abordados por este método? ¿Cómo comprender la mente humana, sus facultades y sus propiedades desde la perspectiva propia de tal método? Una vez la cuestión metodológica quede resuelta, podrán investigarse en adecuación a ese método una serie de temas cardinales, entre los que Evers señala los siguientes: por qué la evolución de las funciones cognitivas produjo seres morales; qué significa ser agente moral; de dónde viene y cómo funciona nuestra predisposición natural para construir sistemas normativos y para producir juicios morales (Evers, 2010: 13). Ya hemos atendido a estas cuestiones en capítulos precedentes (Darwin, 1979; Ayala y Cela Conde, 2001; Ayala, 2006; Joyce; Damasio, 2005; Damasio, 2008; Churchland, 2013; Prinz, 2001; Nichols, 2005; Waal, 1997, 2007; Brosnan y Waal,

²⁸ Evers considera la cognición como algo separado de las emociones, asocia a la cognición procesos tales como el razonamiento, el pensamiento abstracto o la resolución consciente de problemas. No obstante las emociones son un importante proceso cognitivo en la medida en que median entre la información que ofrece la percepción y la respuesta conductual que el individuo da. Por ello considero adecuado incluir las emociones dentro de la cognición. En cualquier caso, para evitar confusión, he de indicar que Evers no expresa explícitamente esta consideración.

2003; Tomasello y Vaish, 2012; Haidt, 2001; Greene;), tendremos ocasión de ver cómo Evers las aborda desde una propuesta enmarcada explícitamente en el ámbito de la neuroética.

El materialismo ilustrado.

La primera de las cuatro conferencias que componen la obra se titula “Cuando la materia se despierta, El espíritu abierto y sus enemigos”. El capítulo comienza tratando algunas cuestiones básicas en neuroética, entre ellas cabe señalar la exposición que hace del advenimiento de la neuroética y la distinción entre ética fundamental y ética aplicada, a la que ya hemos atendido con anterioridad. Esta doble división ha sido posteriormente modificada por la autora y otros investigadores que han colaborado en sus estudios (Evers y Salles, 2017). Estas cuestiones sirven como contextualización e introducción. Tras ello Evers pasa a abordar los temas que componen el verdadero núcleo del capítulo. La autora pretende someter a crítica eso que podemos llamar *enemigos del espíritu abierto*. En primer lugar, lleva a cabo una crítica a la ciencia en la medida en que esta se ha puesto históricamente al servicio de la ideología, de modo afín a las aportaciones de la Teoría Crítica (Habermas, 1999; Apel, 1985; Cortina, 2007). En segundo lugar, centra la crítica en las corrientes científicas que han encarnado lo que ella denomina *psicofobia*, siendo esta una actitud investigadora que desestima el papel de la conciencia y del espíritu en la investigación (resulta de interés aquí atender a: Habermas, 2006). El objetivo de Evers es defender una concepción del cerebro como objeto de estudio de las neurociencias donde este sea tomado como un órgano plástico y dinámico (Jácomo y Serrão, 2013), situando también la conciencia como uno de los objetos para la investigación neurocientífica (Habermas, 2006). Por ello la crítica a la psicofobia será clave. Concluirá el capítulo con una propuesta de *ciencia psicófila* y con un modelo metodológico al que refiere como *materialismo ilustrado*.

Las preocupaciones de Evers en torno al uso ideológico o politizado de la ciencia y a cómo un programa científico debe diseñarse con criterios de responsabilidad serán retomadas con más profundidad más adelante, por ello no profundizaremos ahora en ello. Por su parte, resulta especialmente interesante atender a la crítica que hace a la *ciencia psicófoba* para acercarnos a los conceptos de *ciencia psicófila* y de *materialismo ilustrado*.

La crítica de la autora a la ciencia psicófoba se dirige a dos blancos distintos: el *eliminativismo ingenuo* y el *cognitivism ingenuo* (Evers, 2010: 43-66). El eliminativismo ingenuo, encarnado por la corriente del conductismo (Watson, 1913; Skinner, 1977), presenta una de las formas de psicofobia que la autora denuncia. El conductismo consideraba que solo la conducta, en tanto que acontecimiento público, podía ser tomada como objeto de estudio adecuado para la ciencia empírica de la psicología. Por tanto, toda alusión a la conciencia, así como a los procesos cognitivos y mentales no conscientes²⁹, debía ser eliminada. Es en este sentido en el que el conductismo ha de ser considerado como eliminativista. Cabe especificar que el eliminativismo puede darse en tres formas, éstas son las siguientes: *eliminativismo conceptual*, que aboga por ignorar un término en tanto que esté contaminado por el error, tal sería el caso de los términos referentes a estados mentales, según sugería el conductismo; *eliminativismo contextual*, consistente en eliminar alusiones a ciertos elementos por no considerarlo adecuado en cierto contexto, concretamente el conductismo elimina toda referencia a la conciencia, por no considerarla un objeto de estudio adecuado para las ciencias empíricas; y *eliminativismo ontológico*, que se presenta cuando se niega la existencia de un objeto, por ejemplo de la conciencia o de los procesos cognitivos. No necesariamente han de producirse las tres formas de eliminativismo simultáneamente.

En el contexto del eliminativismo, Evers apunta, además, el caso del *reduccionismo eliminativista*. Es un enunciado aceptado por la ciencia empírica que la conciencia no existe con independencia del cerebro, o bien es una propiedad emergente de la actividad neuronal o bien es una función de esta. Ahora bien, dice la autora, esperar que el estudio empírico del cerebro sea capaz de aportar todo el conocimiento posible acerca de la conciencia es incurrir en reduccionismo eliminativista. Evers reivindica la relevancia de la experiencia subjetiva en el estudio de la conciencia, pues esta es percibida siempre de forma interna y en primera persona. La conciencia no es reducible a meros procesos biológicos, ni tan si quiera en el ámbito de la investigación, ya que presenta una ontología de primera persona. El *materialismo ingenuo* es insostenible por ser ontológicamente reduccionista. Existe aquí una relación con el dualismo de perspectivas metodológicas y epistemológicas que Habermas expone en la sexta conferencia de *Entre naturalismo y religión* (2006), cuyo estudio abordaremos detalladamente en capítulos

²⁹ Evers habla solo de conciencia, pero resulta adecuado incluir también la referencia a los procesos cognitivos y mentales.

posteriores. La perspectiva del partícipe, señala Habermas, está vinculada a un carácter de primera persona, y no es reductible a la perspectiva objetivante del observador. Cuando Evers (2010) dice que la conciencia tiene una ontología de primera persona está poniendo de relieve el hecho de que la experiencia del espíritu no es reductible a la perspectiva del observador, porque es una experiencia constitutivamente subjetiva.

La división interior-exterior es muy problemática, indica la autora. Responde a una división epistemológica respecto a lo que es objetivo, en tanto que observable, y a lo que es subjetivo, en tanto que interno. Sin embargo, en el caso de la conciencia, así como el de la cognición en general, la línea que separa lo interno de lo externo es muy fina. En lo que concierne a la dimensión observable, encontramos la conducta como acontecimiento público, a la que hemos de sumar los datos acerca de la estructura y la actividad cerebral ofrecidos por los estudios neurocientíficos. Así la conciencia queda acotada desde dos flancos de hechos empíricamente observables, la actividad y la arquitectura cerebral, por un lado, y la conducta, por otro. La cognición y la conciencia ocupan el terreno intermedio entre estos hechos observables públicamente, un terreno inaccesible para la observación pública, pero accesible para la introspección y para la experiencia subjetiva e interna. Lo que Evers sugiere es que, si bien el testimonio personal no está nunca libre de los problemas asociados a la subjetividad, es la mejor baza que existe para acercarse al conocimiento de la conciencia como objeto observable, quizás sea la única baza posible, puesto que la conciencia es ontológicamente subjetiva e interna. Por ello se han de considerar como relevantes para la investigación las experiencias subjetivas y los informes personales (Evers, 2010; Habermas, 2006; Conill, 2016).

En otro orden de cosas, el materialismo adopta una posición reduccionista cuando sostiene la posibilidad de hacer una cartografía del pensamiento, además de una cartografía del cerebro. Si bien esta última refiere al mapeado de los procesos neuronales, que diferentes investigaciones están llevando a cabo, aquella refiere al intento de desentrañar el contenido semántico de tales procesos. Dicho de otro modo, una cartografía del contenido semántico de los procesos neuronales permitiría saber qué piensa una persona sin que ella lo diga. Si bien, de acuerdo a las teorías neurosemánticas, la neurociencia quizás llegue a ofrecer información detallada sobre ciertos contenidos de procesos neuronales, parece poco probable que permita facilitar la información completa

de los procesos cognitivos y de la consciencia a un observador en tercera persona, que no experimenta tales procesos ni la consciencia en primera persona.

En cualquier caso, Evers se muestra cauta, y reconoce no estar en posición de dirimir los problemas centrales de este debate. Sin embargo, sí que señala la importancia de los informes personales y de las experiencias subjetivas. A día de hoy, la relevancia de los informes personales sobre la consciencia se debe a que no existe, al menos aún, la posibilidad de tener informes públicos ofrecidos por investigaciones en neurosemántica. Además, en tanto que la consciencia sea considerada como ontológicamente subjetiva, tal y como Evers sugiere, la neurosemántica solo podría ofrecer datos públicamente observables. Desconocemos si es posible inducir al sujeto observador a experimentar subjetivamente estados de consciencia ajenos, correspondientes a determinados procesos neuronales en la arquitectura del cerebro de un individuo concreto y diferente al observador. Esta posibilidad parece inverosímil, pero si llega a ser desarrollada con éxito por la neurociencia, cabría preguntarse si lo que tales sujetos inducidos experimentarían sería un hecho públicamente observable, o si tan solo se estaría facilitando la posibilidad de un nuevo informe personal y subjetivo de unos mismos procesos, esta vez experimentados por nuevos sujetos. En cualquier caso, ante lo problemático e improbable de esta posibilidad, la propuesta de Evers parece adecuada, considerar las experiencias subjetivas como relevantes en relación a la consciencia favorece y facilita el estudio de esta, aunque no sea mediante el estudio de hechos públicamente observables.

Evers reconoce que el materialismo tiene razón cuando considera la consciencia como un fenómeno biológico integrado en el cerebro. Rechaza la dualidad mente-cerebro como dos realidades esencialmente diferentes, siendo una corpórea y la otra inmaterial. Cabe definir el materialismo como una postura monista y naturalista, pero eso no implica que tenga que ser reduccionista o eliminativista. En la actualidad un gran número de neurocientíficos son materialistas sin ser eliminativistas en lo que concierne a la consciencia y a la cognición, ni reduccionistas a la hora de estudiar estas.

La propuesta de Evers establece las siguientes afirmaciones en relación a la consciencia y la cognición, en tanto que objetos de estudio de la neurociencia:

- a) ontológicamente son funciones biológicas resultantes de procesos neuronales.

b) epistemológicamente pueden conocerse mediante la observación de la actividad cerebral y de la conducta, por un lado, y mediante la introspección y los informes personales, por otro.

Todo ello reconfigurando los conceptos de lo físico y lo mental, que deben dejar de referir a realidades distintas, pasando a ser considerados como dos aspectos insoslayablemente unidos.

Por otra parte, Evers dirige su crítica contra la psicofobia hacia el *cognitivismo ingenuo*. Las ciencias cognitivas pusieron de nuevo en primer plano el espíritu, entendido como realidad psicológica y mental. La ingenuidad del cognitivismo que denuncia Evers se debe a la pretensión de estudiar la cognición como una serie de procesos funcionales, en los que la naturaleza física no constriñe ni condiciona el desarrollo de tales procesos. Así, la investigación subestimó la importancia de los elementos cerebrales y de la base biológica sobre la que todos esos procesos se asentaban. Esta cuestión es marcadamente metodológica, la crítica de Evers no hace otra cosa que reivindicar, de nuevo, la necesidad de investigar la cognición y la conciencia de forma conjunta con los aspectos observables, bien conductuales bien de actividad y estructura cerebral, teniendo en cuenta los informes personales.

Hace una segunda crítica al cognitivismo ingenuo. Esta consiste en denunciar que el cognitivismo dejó de lado la importancia de las emociones en su labor investigadora. Cabe señalar que la cuestión de qué función cumplen las emociones en los procesos cognitivos no es cuestión de método, sino de contenido, solo puede saberse qué papel cognitivo juegan las emociones como resultado de la investigación. Frente al predominio del racionalismo en la psicología (Kohlberg, 1992; Piaget, 1987; en este caso cognitiva), Evers reivindica la necesidad de atender a las emociones. En este aspecto su propuesta resulta especialmente relevante para la elaboración de una neuroética de la razón cordial como propuesta metodológica en el campo de la neuroética.³⁰

³⁰ En cualquier caso, cabe hacer algunas aclaraciones en este punto. Evers habla de cognición y de emoción como dos cosas diferenciadas desde el cognitivismo, sin embargo las emociones constituyen un tipo de elementos cognitivo, motivo por el cual considero la cognición como integrada por elementos emocionales. El uso que hago del término *cognición* no es excluyente respecto a las emociones, mientras que el del *cognitivismo ingenuo* sí, según Evers lo expone.

Evers propondrá el *materialismo ilustrado* como el método más adecuado para la neuroética. El materialismo ilustrado se caracteriza por ser un método naturalista, que evita los problemas de la ideología y de la psicofobia en la ciencia, y que es capaz de abordar la conciencia y la cognición como objetos de investigación en las ciencias empíricas. El término fue utilizado originalmente en química y extendido por Changeux (1985, 1993, 2001, 2010) en las neurociencias. El materialismo ilustrado se opone al dualismo y al reduccionismo ingenuo, considera la conciencia como resultado de la evolución y como un producto de la actividad de las estructuras físicas del cerebro, describiendo además el cerebro como un sistema proyectivo, variable, activo y autónomo, en el que las emociones y los valores se incorporan como coerciones necesarias, siendo resultado de una sinergia sociocultural-biológica. El materialismo ilustrado de Evers, además, reconoce la necesidad de complementar los informes sobre observaciones públicas con los informes personales de experiencias subjetivas en torno a la conciencia.

En tanto que método naturalista, considera que todo conocimiento empírico tiene como fuente la naturaleza. Por tanto, las neurociencias, que son una ciencia empírica, han de encontrar en esa naturaleza cualquier objeto que pretenda ser abordado como objeto de estudio. En lo que concierne al cerebro como órgano físico no hay lugar para la discusión, este es un objeto de estudio apropiado para las ciencias empíricas, un órgano físico resultado de un proceso de evolución biológica, que realiza unos procesos y una actividad observables. La neurociencia, cabe señalar, no solo pretende estudiar el cerebro, sino estudiarlo en relación con la conducta, los procesos cognitivos y la conciencia. En tanto que el materialismo ilustrado pone las condiciones necesarias para que una investigación de este tipo sea posible será adecuado como método científico. Además, sostiene Evers, el materialismo ilustrado supone un marco a partir del cual la reflexión filosófica en torno a las cuestiones neurocientíficas será adecuada y productiva, permitiendo una neuroética verdaderamente interdisciplinar, en la que ciencias empíricas y filosofía moral trabajen conjuntamente. Evers considera que el materialismo ilustrado encarna una propuesta metodológica adecuada para la neuroética, ya que permite establecer un marco científico adecuado y filosóficamente fructífero.

El materialismo ilustrado es definido por Evers presentando las siguientes características (Evers, 2010: 71):

1. La conciencia es un objeto apropiado para la investigación científica. El materialismo ilustrado adopta una concepción evolucionista de la conciencia, considerando esta como una parte irreductible de la realidad biológica, siendo una función del cerebro aparecida en el curso de la filogénesis de las especies.
2. En relación al estudio de la experiencia consciente y subjetiva, el materialismo ilustrado considera relevantes para la investigación tanto la información objetiva, resultado de observaciones y medidas anatómicas y fisiológicas, como la subjetiva, obtenida por introspección y en primera persona.
3. Describe el cerebro como un órgano plástico, proyectivo y narrativo, resultado de una simbiosis o una sinergia sociocultural-biológica. Este puede actuar consciente e inconscientemente de manera autónoma.
4. La emoción es considerada como la marca distintiva de la conciencia, siendo gracias a ella que la materia puede despertar y producir un espíritu dinámico, flexible y abierto.

Si sometemos a análisis estas características del materialismo ilustrado podremos apreciar que no todas ellas son cuestiones referentes al método, sino que el materialismo ilustrado adopta como punto de partida algunas ideas que son resultado, y no condiciones iniciales, de la investigación. Los cuatro puntos anteriores pueden dividirse en dos grupos. El primero de ellos estaría integrado por los puntos 1 y 2, y es el que propiamente atiende a cuestiones de método. El punto 1 establece la pertinencia de considerar la conciencia y la cognición como objeto de estudio científico. El punto 2 expone qué datos son relevantes en la investigación. En cuanto a los puntos 3 y 4, estos no corresponden a cuestiones metodológicas, sino que suscriben una serie de tesis acerca del cerebro y la conciencia que precisan atender a una investigación empírica que las respalde. Así pues, estos dos últimos puntos no exponen características propias del método del *materialismo ilustrado*, sino tesis que pueden ser respaldadas gracias a la aplicación del método en investigaciones concretas.

Podemos ofrecer una definición del materialismo ilustrado más precisa, que atienda únicamente a las cuestiones de método, atendiendo a los dos primeros puntos de la serie anterior, y que además añada algunos otros, considerando como objeto de estudio no solo la conciencia, sino la cognición en general. Esta definición, más precisa, presentaría las siguientes características:

1) *Paradigma naturalista*: el materialismo ilustrado es un método de investigación científica que opera en el paradigma naturalista y cuyos objetos de estudio han de ser objetos naturales.

2) En lo que concierne a la conciencia y a la cognición:

- a) *Objeto de investigación*: conciencia y cognición son objetos apropiados de estudio para la investigación científica. Ambas son realidades naturales, resultado de la evolución y parte irreductible de la realidad biológica, con el cerebro como soporte físico.
- b) *Información relevante*: tanto la información objetiva, resultado de observaciones conductuales y mediciones anatómicas y fisiológicas, como la subjetiva, obtenida por introspección, son relevantes para la investigación, puesto que ambas ofrecen datos en torno a esos objetos de estudio que son la conciencia y la cognición. La perspectiva del partícipe en primera persona se vuelve relevante en la medida en que la observación de la conciencia no puede hacerse más que subjetiva e internamente, por ser constitutivamente subjetiva.
- c) *Monismo ontológico*: lo mental y lo físico se consideran insoslayablemente entrelazados.
- d) El materialismo ilustrado es no eliminativista, tanto epistemológica como ontológicamente, y no reduccionista.

Como puede verse, el primer punto clarifica qué tipo de método es el materialismo ilustrado, en qué paradigma se sitúa y qué objetos de estudio corresponden a sus competencias investigadoras. El segundo punto compila los dos primeros puntos de la definición dada por Evers, expuesta más arriba, ambos referentes a la conciencia en tanto que objeto de estudio, a los cuales hemos añadido dos puntos más, de acuerdo a las ideas que la propia autora expone en su obra.

Una vez establecidas las características básicas del método de investigación, cabe atender a las posibilidades que este método presenta. Evers señala muy acertadamente que viejos problemas filosóficos emergen ahora en el seno de las neurociencias, cuestiones como la de la voluntad libre, la responsabilidad, la relación entre emociones y conocimiento, la de qué es la agencia moral o la de cómo funciona la producción de

juicios morales, adquieren máxima actualidad en el estudio neurocientífico. Como bien indica Evers, el foso explicativo entre las neurociencias y la conducta moral es más estrecho que, por ejemplo, el que esta guarda con la genética. Dicho de otro modo, las neurociencias pueden ofrecer una explicación, de forma más eficiente que la genética, acerca de por qué pensamos, sentimos y nos comportamos moralmente del modo en que lo hacemos. El estudio del cerebro lleva a cabo uno de los acercamientos más pronunciados hasta la fecha entre las causas de la conducta moral y la propia conducta. Ofrece, a este respecto, datos muy valiosos para la comprensión de la toma de decisiones, de la elección, de la adquisición de un carácter o de disposiciones morales, de la producción de juicios morales, y de nuestras nociones de bien y de justicia, entre otras cuestiones.

Las neurociencias, para las que Evers propone como método el *materialismo ilustrado*, son capaces de ofrecer información de este tipo acerca de sus objetos de estudio. Sin embargo, ha de señalarse que toda esta información es de carácter descriptivo y explicativo, versa sobre cómo y por qué la conducta se produce en los términos en que se produce. Sin embargo, en cuanto a la ética, esta no versa solo sobre cómo la conducta moral es y por qué, sino también sobre cómo debe ser y por qué debe ser así (Apel, 1985; Habermas, 2010; Cortina, 2007, 2009). Por tanto, la labor del materialismo ilustrado, en tanto que método naturalista, queda limitada a la vertiente descriptiva y explicativa de la neuroética. Evers es consciente de ello, por eso asegura que el materialismo ilustrado es inmune a la crítica del reduccionismo naturalista, como tendremos ocasión de ver. Lo que Evers plantea no es un método que proponga deducir un *deber moral* a partir de una descripción y explicación científica de lo que *es*. Lo que propone es un método que ofrezca un marco adecuado para la investigación científica de tal modo que la reflexión filosófica pueda ser compatible y ejercerse adecuadamente, permitiendo que ciencias empíricas y filosofía moral operen conjuntamente en el ámbito de la neuroética. Evers reconoce que la tarea de la ética no es reductible al paradigma naturalista ni al método científico, y no pretende hacer tal cosa mediante la propuesta metodológica que encarna el *materialismo ilustrado*. La propia autora señala en algunos puntos la insuficiencia del paradigma naturalista en lo que concierne a la fundamentación de principios, normas y valores morales. Advierte, por ello, del error lógico y falaz consistente en confundir lo que “es” con lo que “debe ser” (Evers, 2010: 144-148). Las competencias del materialismo ilustrado, en tanto que método naturalista, y las de la ética, cuyo método es

filosófico —aun si atiende a datos científicos en su reflexión (Moore, 1983; Teehan y diCarlo 2004)—, son diferentes y han de trabajar de forma complementaria para un adecuado desarrollo de la neuroética (Buchanan, 2009; Apel, 1985; Habermas, 2006; Aranguren, 1994; Gould, 1999; Hume 1988, 1953; Moore, 1983; Salles, 2013, 2016, 2018; Evers, 2010; Cartwright, 2010; Teehan y diCarlo, 2004; García-Marzá, 2016; Conill, 2012).

Evers considera que la naturaleza humana está compuesta por la arquitectura cerebral y por la influencia del entorno social y cultural (Ayala y Cela Conde, 2001; Hauser, 2008; Evers, 2010; Aranguren, 1994; Piaget, 1987; Kohlberg, 1992; Cortina, 2012a). En sintonía con esa idea de naturaleza, concibe la neuroética como una disciplina compuesta por el estudio empírico, con el *materialismo ilustrado* como método, y por la labor filosófica, que se ocupa de reflexionar sobre qué normas morales deben guiar la conducta humana moralmente, en el plano social y cultural (Apel, 1985; Habermas; Cortina; Korsgaard, 2000). Mientras que la vertiente empírica ofrece datos descriptivos orientados a la explicación y la predicción, la vertiente filosófica se ocupa de los criterios de justificación.

Comparto con Evers la tesis de que esta división de competencias entre neurociencia y filosofía moral es la adecuada. La conducta moral cuenta con una base biológica y neuronal que la condiciona, esta compone un hecho de la naturaleza, del cual no puede derivarse un deber moral sin incurrir en reduccionismo naturalista. Las estructuras morales biológicas componen una serie de hechos cuyo estudio es científicamente abordable. Por su parte, los contenidos morales que se asientan sobre esas estructuras son contingentes y culturales, siendo su fundamentación normativa una propia de la filosofía moral.

Neuroeducación, epigénesis proactiva y responsabilidad naturalista.

En el tercer capítulo de su libro, “La base neural de la moralidad, La pertinencia normativa de las neurociencias”, Evers atiende a la relación entre las neurociencias y el establecimiento de la normatividad moral. Para ello lleva a cabo un análisis de las aportaciones que la ciencia hace en lo que concierne a la base neuronal del comportamiento moral, indicando también qué límites encuentra. Delimita así las

competencias de la ciencia, reconociendo a la filosofía la competencia de investigar en torno a las cuestiones para las cuales el método naturalista es insuficiente.

Dice Evers que el ser humano presenta predisposiciones biológicas para desarrollar sistemas de valores complejos y variados, presentando cada uno de nosotros unas tendencias preferenciales y valorativas que pueden diferir en distintos individuos (como ya vimos con: Ayala y Cela Conde, 2001; Aranguren, 1994; Hauser, 2008). Los valores, la emisión de juicios morales, el asentimiento o rechazo a ciertas normas, la obligación moral, etc., no cumplen solo funciones prácticas, sino que además se imponen necesariamente desde nuestra biología, ya que somos estructuralmente morales (Aranguren, 1994; Cortina, 2012a). Somos animales sociales con una cognición compleja marcada biológicamente, a los que la normatividad moral se les presenta como un problema (Korsgaard, 2000; Darwin, 1979; Ayala y Cela Conde, 2001; Ayala, 2006; Joyce; Damasio, 2005; Damasio, 2008; Churchland, 2013; Prinz, 2001; Nichols, 2005; Waal, 1997, 2007; Brosnan y Waal, 2003; Tomasello y Vaish, 2012). La biología nos hace constitutivamente morales, sin dotarnos de unos códigos morales en forma de contenidos, estos son contingentes e históricos, correspondiendo a la filosofía moral ofrecer criterios para su fundamentación ética.

Evers señala cuatro tendencias preferenciales innatas e íntimamente relacionadas surgidas en el curso de la evolución de nuestra especie. Estas son el interés por uno mismo, el deseo de control y seguridad, la disociación de lo que es considerado como desagradable o como una amenaza y el sentimiento de simpatía —entendido como empatía positiva—.

En relación a ello, apunta dos cuestiones relacionadas con la pertinencia normativa de las neurociencias. La primera de ellas refiere a la situación egocéntrica y endogámica en la que el ser humano se halla inmerso dadas sus predisposiciones biológicas y neuronales. Por un lado, la simpatía se orienta egocéntrica y endogámicamente dado que la evolución así lo ha dispuesto (Darwin, 1979; Ayala y Cela Conde, 2001; Ayala, 2006; Damasio, 2008; Churchland, 2013; Waal, 1997, 2007; Brosnan y Waal, 2003; Tomasello y Vaish, 2012; Haidt, 2001; Greene; Cortina, 2009, 2012a, 2017). Por otro, nuestro cerebro nos sitúa en un egocentrismo cerebral a nivel psicológico, somático y espacio-temporal. Estos hechos nos inclinan hacia una miopía tanto biológica como cultural a la hora de ejercer nuestra agencia moral. La pertinencia normativa de las neurociencias en

este punto consiste en hacer un diagnóstico de la difícil situación neurobiológica en la que nos encontramos si pretendemos ser actores morales que defiendan y practiquen una ética de corte universal.

La segunda de las cuestiones que Evers apunta, en relación a la pertinencia normativa de las neurociencias, permite vislumbrar una posible solución a los problemas que plantea ese carácter egocéntrico y endogámico al que estamos predispuestos. Dado que el cerebro es un órgano plástico y maleable que se desarrolla en interacción con el medio, cabe la posibilidad de orientar epigenéticamente el desarrollo tanto de contenidos como de estructuras morales, de forma que se puedan vencer, controlar y/o guiar las predisposiciones biológicas y conductuales que nos resulten moralmente problemáticas. Dado que la relación entre nuestro cerebro y nuestras sociedades es recíproca y simbiótica, influyéndose mutuamente, tenemos el poder de ejercer una mejora epigenética de las disposiciones cerebrales, lo cual supondrá una mejora de las sociedades en las que vivimos, y viceversa.

Es posible que escojamos utilizar bien nuestro poder y también evolucionar, tanto biológica como culturalmente, para convertirnos en lo que consideramos criaturas “mejores” que desarrollan sociedades más avanzadas. (Evers, 2010: 190)

Neuroeducación moral.

En este sentido, la educación ocupa un lugar central como vía para orientar el desarrollo epigenético del cerebro (Evers, 2010; Pallarés-Domínguez y Richart, 2018). La educación es un proceso complejo, mediante el cual se produce una transmisión de conocimientos, costumbres, valores y modos de actuar entre individuos en el seno de una sociedad o grupo. Estos rasgos, adquiridos mediante la educación, son históricos, culturales y contingentes, y se adoptan de forma singular y específica en cada individuo, en interacción con diferentes miembros, grupos, comunidades e incluso culturas que integran la sociedad en la que el individuo vive (Ayala y Cela Conde, 2001). Estos contenidos culturales precisan una estructura psicobiológica sobre la que asentarse, un soporte neuronal que haga al sujeto susceptible de ser educado (Aranguren, 1994; Hauser, 2008; Cortina, 2012a, 2017). El cerebro es ese soporte, pone las condiciones de posibilidad para la educación.

La educación puede ser entendida, pues, en dos estratos, el de estructura y el de contenido. Las estructuras suponen la base fisiológica que presenta el educando desde su nacimiento y que está en continuo desarrollo en interacción con el entorno, desde el cual puede ejercerse la educación, de modo que el desarrollo de estas estructuras y los contenidos que se alojan en ellas vayan siendo dirigidos por los educadores, en sintonía con los objetivos educativos que se persigan. La educación moral, por tanto, puede definirse desde una perspectiva neurocientífica como un proceso dirigido mediante el cual se guía el desarrollo epigenético de las estructuras morales durante la ontogénesis, posibilitando la adquisición de rasgos culturales considerados como contenidos, incluyendo en estos el *ethos* o carácter moral, así como diversas competencias y habilidades sociales y morales.

Plasticidad cerebral y epigénesis proactiva.

Las estructuras cerebrales se ven sujetas a continuo cambio por el proceso de desarrollo neuronal llamado epigénesis. La epigénesis se define como el proceso en el que un sujeto se desarrolla en su interacción con el medio. Según explica Evers, en relación al modelo de Changeux, la epigénesis refiere a la selección de sinapsis cerebrales establecidas en función de su actividad, siendo ésta una forma de impronta cultural (Evers, 2010: 137; Salles, 2017). El desarrollo del cerebro presenta una gran versatilidad, puesto que el 70% de este se produce de forma epigenética (Cortina, 2012a: 220). La plasticidad neuronal, neural o sináptica, también llamada neuroplasticidad, es una propiedad del funcionamiento neuronal en el proceso de conexión y comunicación con el resto de neuronas. Cada una de estas presenta unas 50.000 conexiones de promedio, lo cual sumado al número de neuronas que componen nuestro cerebro hace de este un órgano con un alto nivel de plasticidad (Evers, 2010: 138; Jácomo y Serrão, 2013)³¹.

Estos datos son muy relevantes para la educación moral. La plasticidad neuronal y el desarrollo epigenético que el cerebro experimenta ofrecen a la educación la posibilidad de guiar el desarrollo de las estructuras cerebrales, también en lo que concierne a la dimensión moral. La estructura cerebral se configurará de diferente modo en función de cómo los elementos del entorno se presenten al sujeto, de modo que tanto los educadores como la sociedad tienen la posibilidad y la responsabilidad de ejercer una

³¹ Para ver más sobre la *variabilidad* en las funciones y el desarrollo, así como sobre la *epigénesis neuronal*, cabe atender al modelo de Changeux expuesto por Kathinka Evers (2010).

educación y una socialización que moldee de la forma más adecuada las estructuras cerebrales del educando.

En relación a ello, resulta de especial interés la propuesta de Evers. Esta plantea la posibilidad de ser proactivos desde un punto de vista epigenético, ya que las estructuras cerebrales se moldean y desarrollan en gran medida en la interacción con el entorno se puede tomar parte activa en la construcción de tales estructuras. Eso significa que los valores y normas de una sociedad o cultura pueden componer en sus miembros una serie de estructuras cerebrales ajustadas a los mismos en un alto grado (Evers, 2010: 136-144; Evers, 2015). De este modo la predisposición genética, que condiciona la acción moral endogámicamente (Waal, Tomasello, Ayala), puede compensarse mediante la disposición epigenética, que puede inculcar valores de corte universal y de justicia desinteresada. La acción proactiva en la epigénesis, a través de la educación y los procesos de endoculturación, interrelaciona el elemento estructural de la moral con los contenidos, habilidades y carácter que el educando va adquiriendo epigenéticamente.

La arquitectura de nuestros cerebros determina quiénes somos y el tipo de sociedades que desarrollamos; pero nuestras estructuras sociales también tienen un fuerte impacto sobre la arquitectura del cerebro, en particular a través de la huella cultural que está allí epigenéticamente almacenada. El camino, por consiguiente, está abierto para que seamos *proactivos desde el punto de vista epigenético*. (Evers, 2010: 136)

Además, plantea Evers, según parecen indicar recientes estudios genéticos, cabría la posibilidad de que algunos caracteres adquiridos epigenéticamente a nivel cerebral pudiesen dejar una impronta genética. En sentido lamarckista podríamos decir que, si esto es cierto, hay caracteres adquiridos que pueden llegar a heredarse. En el plano de la educación moral esta cuestión es especialmente relevante, si durante generaciones se educa en valores que consideramos moralmente deseables, como por ejemplo la paz o la cordialidad, estos pueden generar una impronta genética que podría llegar a hacer que exista cierta tendencia hacia este tipo de caracteres de forma innata. Igual sucede si durante generaciones se educa en valores que consideramos reprobables, como la violencia y el odio.

Las teorías sobre la epigénesis y la sinapsis de las improntas culturales y sociales en nuestra arquitectura cerebral abren la puerta a ser epigenéticamente proactivo, lo que significa que podemos influir culturalmente en la organización de nuestro cerebro con el objetivo del

autoperfeccionamiento, individualmente como socialmente, y cambiar nuestras predisposiciones a través de un mejor ajuste de nuestro cerebro a las culturas y estructuras sociales. (Evers, 2015)

Es una responsabilidad para el educador, y para la educación en general, así como para la sociedad, favorecer que el educando reciba una educación adecuada, en aras de que si se produce una impronta genética de algún tipo sea beneficiosa para las futuras generaciones. Las condiciones en que esa impronta puede producirse, en caso de que se produzca, así como el número de generaciones necesarias para que se manifieste de forma significativa no han sido aún determinados por los genetistas (Salles, 2017).

Cabe preguntarse qué papel juega la educación en el *ethos*, las habilidades y las competencias sociales y morales del educando. A nivel genético, biológico y neuronal los individuos pueden presentar predisposiciones que faciliten o dificulten la adquisición de ciertas habilidades o cierto carácter, no obstante la plasticidad neuronal ofrece la posibilidad de enfrentarse a esas predisposiciones (Jácomo y Serrão, 2013). En este sentido la educación será clave. Al forjar un carácter o *ethos* o al generar un hábito se están produciendo disposiciones estructurales, se genera una segunda naturaleza por apropiación (Aranguren, 1994). Por epigénesis proactiva podría producirse también, en este nivel, una impronta genética.

Por último, cabe atender a la adquisición de los diferentes sistemas de normas. Éstos son, como ya hemos indicado, contingentes y culturales, y precisan de una estructura sobre la que asentarse. Estos sistemas presentan diferentes constelaciones de valores y de normas que tratan de orientar la conducta, pudiendo apelar a diferentes elementos a la hora de ser justificados o fundamentados éticamente.

Prosocialidad y predisposiciones biológicas.

La evolución ha hecho del ser humano un animal complejo en lo que concierne a su dimensión social y moral. Por un lado, encontramos unos códigos de conducta social grabados en nuestra biología que predisponen al cuidado de los más cercanos, sentando las bases de una conducta social y moral orientada endogámicamente (Waal, 1997, 2007; Tomasello, 2013; Ayala, 2006; Ayala y Cela Conde, 2001; Greene, 2013; Churchland, 2013; Evers, 2010; Cortina, 2009, 2012a, 2017). Por otro lado, encontramos una serie de valores y normas morales adquiridos culturalmente, que en muchas sociedades tienen la pretensión de ser universalizables y reconocen el valor de toda vida humana como digna

(Habermas, 2010; Koorsgard, 2000; Cortina, 2007, 2009). Se genera así una disonancia entre lo que la biología marca y lo que ciertos contenidos morales sugieren. Este hecho supone un reto en muchos sentidos, del cual la educación no queda exenta. En lo que a ella concierne, cabe preguntarse qué papel vamos a reconocer a esos códigos biológicos de corte endogámico y cuál a esos otros códigos culturales universalistas (Cortina, 2012a: 221-222; Cortina, 2017).

Para evitar errores lógicos que conduzcan a reduccionismos de tipo naturalista no debemos considerar la existencia de ciertos códigos biológicamente predispositivos como un criterio o una razón para realizar prescripciones morales. La existencia de tales códigos y un conocimiento riguroso de estos suponen, sin duda, datos valiosos que han de ser tenidos en cuenta a la hora de educar moralmente (Moore, 1983; Teehan y diCarlo 2004), pero no pueden constituir por sí mismos una propuesta de deber moral, para ello se necesita que la filosofía moral los respalde, fundamentándolos adecuadamente, en los casos en que sea posible (Cortina, 2012a; Habermas, 2006, 2010; Apel, 1985, Korsgaard, 2000; Haidt, 2001).

Esta disonancia entre unos y otros códigos, biológicos y morales, pone de relieve que hemos de preguntarnos no por lo que es, sino por lo que debe ser.

La posibilidad de cultivar el cerebro deja en nuestras manos llevar a cabo esas acciones autoformativas, de las que hablaba Kane, por las que vamos forjando nuestro carácter. En las situaciones en las que entran en competencia distintas redes neuronales y hemos hecho el esfuerzo de optar por uno de los *inputs* vamos forjándonos un carácter que va modificando también nuestro cerebro. *Hacia dónde* queremos forjarlo es la pregunta que no contesta ninguna descripción del cerebro, porque *no es la pregunta por lo que hay, sino por lo que debe haber*. (Cortina, 2012: 223)

La posibilidad de forjar el propio desarrollo neuronal refiere a la cuestión de la plasticidad neuronal y la epigénesis (Evers, 2010, 2015; Salles, 2017; Jácomo y Serrão, 2013). Además, Cortina señala la incapacidad de las ciencias para abordar desde el paradigma naturalista la cuestión de la fundamentación ética (Cortina, 2007). La educación moral, por tanto, ha de atender tanto a la filosofía moral como a la ciencia, pero sin confundir las competencias de cada una de ellas (Cortina, 2012a).

Puesto que el ser humano tiene arraigada en su naturaleza una inclinación a actuar egoísta y endogámicamente, la tarea de la educación moral debe dirigirse a modificar epigenéticamente esa predisposición. No puede lograrse al completo, ni se pretende, que

los individuos se despojen de esa dimensión biológica, emocional y predispositiva de carácter endogámico y egoísta. De lo que se trata es de formar sujetos capaces de reconocer que esas inclinaciones y preferencias no responden a razones morales, sino a predisposiciones biológicas, siendo capaces de ser críticos con tales códigos, rechazándolos cuando no puedan ser respaldados racionalmente (Haidt, 2001; Cortina, 2012a). Esto permitirá al educando adoptar criterios de conducta que respondan no solo a elementos emocionales y a predisposiciones biológicas, sino también a valores fundamentados desde la razón, pudiendo reconocer su validez en función de un análisis crítico. De este modo, la educación moral podrá hacer que el educando sea capaz de compatibilizar sus tendencias biológicas y emocionales con una postura ética fundamentada desde la razón (Evers, 2010: 136-144), sin desatender por ello el papel de los elementos cordiales siempre presentes en el ejercicio de la razón (Cortina, 2009, 2017).

Neuroeducación moral desde la razón cordial: una educación moral emotivo-racional.

Las neurociencias han puesto de relieve el importante papel de las emociones en el ejercicio de diversos procesos cerebrales y facultades psicológicas (Salles, 2015; Damasio, 2008, 2009; Eysenck, 2000; Greene; Haidt, 2001, 2013; Le Douarin, 1999; Morgado, 2010; Churchland, 2013). Desde la aparición de las emociones hace cientos de millones de años estas se han ocupado de procesar los elementos percibidos del entorno y ofrecer una respuesta adecuada a tales elementos. El desarrollo del neocórtex, que trajo formas complejas de pensamiento racional, abstracto y metacognitivo tuvo lugar en la evolución humana hace solo unos cientos de miles de años. Las funciones que el neocórtex permite tienen a menudo un papel complementario al de las emociones, sin sustituirlo (Damasio, 2008, 2009; Ayala, 2006; Ayala y Celia Conde, 2001; Churchland, 2013; Darwin, 1979; Forbes y Grafman, 2010; Goldberg, 2002; Greene; Haidt, 2001, 2013; Tomasello, 2013; Waal, 1997, 2007). La razón humana, asociada al desarrollo del neocórtex, y aun habiendo aparecido muy posteriormente a las emociones, permite, en algunos casos, transformarlas, gobernarlas, revisarlas, modificarlas y sustituirlas.

La neuroeducación, pues, ha de atender a este carácter emotivo-racional del ser humano. La posibilidad de educar las emociones ha de ser aprovechada para limar asperezas entre las tendencias conductuales que marca la biología, por un lado, y los

dictados de la ética, por otro (Evers, 2010: 125-133; Cortina, 2009, 2012a, 2017; Salles, 2015). Para ello se ha de trabajar en una doble vía. En primer lugar se han de educar las emociones de forma que el sentimiento moral se extienda, si bien presenta de forma innata un carácter endogámico se puede inculcar una tendencia universal. En segundo lugar, se ha de educar la razón para que las intuiciones emocionales sobre lo que está bien y mal puedan ser sometidas a crítica adecuadamente, revisándolas, corrigiéndolas y descartándolas cuando sea necesario (Haidt, 2001, 2013; Cortina, 2007, 2009, 2012a; González Esteban, 2013; Reverter Bañón y Medina-Vicent, 2018; Cayuela, A., 2012). Una vez hecho esto, el educando se habrá convertido en un agente moral crítico y responsable, con una dimensión moral marcada por las emociones pero supervisada racionalmente.

Como ya hemos visto, la neurociencia pone de relieve algunos aspectos que son cruciales para la educación. Existen una serie de predisposiciones biológicas que dirigen la conducta social y moral de forma egocéntrica y endogámica. Sin embargo, el cerebro se desarrolla en un 70%, aproximadamente, en interacción con el medio, esto permite, por un lado, reforzar aquellas tendencias que consideremos valiosas, por ejemplo las que favorecen la prosocialidad, y, por otro, debilitar aquellas que fomenten conductas moralmente reprobables, como la agresividad. Existe, además, la posibilidad de generar una impronta genética de aquellos caracteres adquiridos, por epigénesis proactiva (Evers, 2010, 2015). La neurociencia pone de relieve la estrecha relación entre razones y emociones, por tanto, habrá que educar teniendo en cuenta este hecho, para así favorecer que el educando logre compatibilizar su dimensión emocional con un análisis crítico de los códigos que rigen su conducta. De este modo será posible limar asperezas y estrechar la distancia que hay entre los códigos que nuestra naturaleza ha inscrito en nosotros y los códigos que la ética y la razón señalan como los moralmente más deseables.

La pertinencia normativa de las neurociencias como responsabilidad naturalista.

Evers considera que existe un sentido en el que las neurociencias resultan normativamente relevantes. La pertinencia normativa de las neurociencias consiste, según la autora, en hacer un diagnóstico de la situación neural del ser humano, que es egocéntrica y endogámica, y ofrecer conocimientos acerca del cerebro que permitan contrarrestar los problemas derivados de esa situación, dado el carácter variable y plástico

del cerebro (Jácomo y Serrão, 2013). Ser proactivos epigenéticamente supone, así, una responsabilidad que presenta ciertos aspectos normativos y naturalistas al mismo tiempo. En la medida en que las neurociencias nos dicen cómo podemos construir nuestras estructuras cerebrales, incluso con una huella genética que sería hereditaria, podemos transformar la naturaleza humana a partir del ejercicio y la aplicación de normas, principios y valores morales. De este modo, las recomendaciones de la ética, fundamentadas desde la filosofía, influirían en la generación de estructuras cerebrales, que transformarían nuestra naturaleza, haciéndola afín a las prescripciones morales ofrecidas por la ética.

Estas consideraciones no incurren en un reduccionismo naturalista, dice Evers. La autora reconoce acertadamente que la descripción de los procesos neuronales subyacentes a las conductas morales no puede más que explicar por qué y cómo estas se producen a nivel neural. La explicación, en cualquier caso, no implica una justificación, la distinción entre hecho y valor, entre ser y deber ser, ha de ser mantenida en las neurociencias.

Si admitimos que los descubrimientos neurocientíficos tienen una pertinencia teórica y metodológica para la ética, y que las estructuras socioculturales y neuronales se desarrollan en simbiosis, y son causalmente pertinentes unas para las otras, de esto se deduce que las neurociencias tienen igualmente una pertinencia normativa, pero no en el sentido de que enunciarían hechos que implican o constituyen normas. Los datos neurocientíficos pueden entregar un conocimiento de una importancia fundamental acerca de la moralidad humana sin implicar normas o recomendaciones. A menos que se cometa lo que se conoce en filosofía con el nombre de *sofisma naturalista*, los juicios normativos (éticos u otros) no pueden ser lógicamente derivados de enunciados descriptivos, por ejemplo de enunciados acerca de datos científicos sobre la naturaleza neurobiológica de este tipo de juicios. En virtud del hecho de que es una objeción estándar en oposición a las tentativas naturalistas para unir los hechos y los valores, este argumento es digno de atención. No obstante afirmaré que, en cuanto objeción opuesta al materialismo ilustrado, está fuera de lugar y no presenta ningún interés para una neuroética que se inscribe en ese marco teórico. (Evers, 2010: 143-144)

Evers afirma, así, que el *materialismo ilustrado* no incurre en reduccionismo naturalista, pues este método se ocupa solo de atender a los hechos desde el prisma científico, no a proponer normas o prescripciones morales atendiendo a esos hechos. Es decir, el materialismo ilustrado evita el reduccionismo naturalista en tanto que es propuesto como un método adecuado para una neuroética descriptiva y explicativa, de corte científico, no para una neuroética fundamental, en sentido ético. De este modo el

materialismo ilustrado permite complementar el ejercicio de la ciencia y de la filosofía moral en la neuroética, sin incurrir en reduccionismo naturalista.

A pesar de mantener esta distinción entre hechos y valores, Evers apunta que sí existe una relación entre ellos, no de justificación o de implicación, sino de lazos empíricos, especialmente causales. Todos los sistemas de normas, las normas que los integran y los diferentes individuos en los que estos se actualizan se hallan integrados en un proceso histórico, social y biológico que ha conducido al surgimiento y asentamiento de tales elementos. Una tarea central de la neuroética fundamental, asegura, consiste en investigar y descubrir la red de conexiones causales entre la neurobiología, las sociedades, las culturas y la historia, que propiciaron la enunciación de ciertas normas y el asentimiento ante estas en un momento dado. Esta tarea permitirá dirimir la naturaleza de estas normas y los motivos del asentimiento que les ha otorgado vigencia y validez, pudiendo así analizar si existen o no razones para asentir a la norma, una postura muy afín a las propuestas de las éticas dialógicas en cuya tradición se inscribe la presente investigación.

El cuarto y último capítulo de la obra, titulado “La responsabilidad naturalista, Hacia una filosofía para la neuroética”, presenta una propuesta que tiene un fuerte calado filosófico. El capítulo comienza retomando la cuestión de la construcción de la naturaleza humana, en tanto que se puede actuar sobre el desarrollo de las estructuras cerebrales siendo proactivos epigenéticamente (Evers, 2010, 2015; Salles, 2017). El carácter constructivo del ser humano en tanto que criatura neurocultural nos otorga, dice Evers, un poder sobre aquellos aspectos de nuestra vida que dependen o que están vinculados con esa dimensión constructiva. Un desarrollo epigenético adecuadamente dirigido nos permite adquirir un poder sobre el sentido de nuestras vidas y sobre el control de las emociones. Evers atiende en este punto, como foco de reflexión, a aquellos aspectos concernientes la construcción de normas.

Para ello recupera y expone aquí algunas propuestas que ya ha esbozado en el tercer capítulo. Entre ellas encontramos la necesidad de superar los problemas que vienen asociados al carácter egocéntrico y endogámico del ser humano, propiciado por su constitución biológica y neurológica, así como la posibilidad de superar esas dificultades mediante un desarrollo epigenético dirigido. Por último, retoma, y esta es la cuestión en la que nos interesa incidir, su idea de una responsabilidad naturalista.

La responsabilidad naturalista, que se manifiesta igualmente en la neuroética fundamental, implica descifrar la red de las conexiones causales que existen entre las perspectivas neurobiológicas, socioculturales e históricas según las cuales una norma moral se encuentra enunciada en un momento dado de la historia humana. (Evers, 2010: 166)

La responsabilidad naturalista debe descifrar cuáles son las condiciones neuronales, sociales e históricas en las que se produce dicho asentimiento, para poder así sopesar las condiciones en las que se produce el procedimiento de establecimiento de lo normativo. Ulteriormente, podrá también distinguirse el carácter universal de las estructuras que subyacen a esa toma de decisión consensuada que aprueba la norma, así como el carácter local de aquellas normas constituidas de forma contingente en el seno de cada sociedad.

La búsqueda de universalidad normativa, en un sentido ético, va más allá de la universalidad de las estructuras y procesos subyacentes a los procedimientos de asentimiento ante normas. El materialismo ilustrado, en el sentido en que lo describe Evers, no está capacitado para abordar qué contenidos morales o que procedimientos de establecimiento normativo son los adecuados, pero sí insta un marco metodológico y teórico a partir del cual la filosofía puede atender a estas cuestiones a la luz de las neurociencias. Evers aboga por la búsqueda de esa universalidad normativa mediante criterios racionales, en oposición a aquellos criterios de carácter dogmático. La búsqueda de una base racional para el universalismo normativo es una idea heredera del proyecto de educación ilustrada, con el que Evers simpatiza. En esta línea de pensamiento podemos encuadrar los procedimentalismos normativos (Habermas; Apel; Rawls; Cortina). Es necesario contar con procedimientos que posibiliten un alejamiento de esa naturaleza neurobiológica orientada egocéntrica y endogámicamente, que permitan el establecimiento de normas de forma universal, imparcial y desinteresada, evitando así desestimar los intereses de aquellos individuos que resultan lejanos para el individuo, dada su distancia cultural, espacial o temporal.

Cuando los xenófobos empáticos aplican programas normativos (ya sean políticos, éticos u otros), es poco probable que tengan en cuenta espontáneamente a sus congéneres más alejados [...]. La concepción egocéntrica tiene componentes al mismo tiempo biológicos y culturales. Estos últimos deben ser correctamente comprendidos con el objeto de que la perspectiva normativa pueda ser extendida para incluir lo que está alejado en el tiempo, en el espacio o desde un punto de vista cultural [...]. (Evers, 2010: 172-173)

Los procedimentalismos que buscan el establecimiento de normas mediante criterios de imparcialidad y racionalidad (Apel; Habermas; Cortina; Rawls) suponen propuestas adecuadas para resolver los problemas asociados al egocentrismo y a la endogamia biológica y neural. Estas propuestas ofrecen mecanismos y procedimientos adecuados para establecer, revisar y corregir lo normativo. Por todo ello, cobra fuerza la propuesta de una neuroética de la razón cordial, como método adecuado para el correcto ejercicio de la neuroética.

Capítulo 6. Reduccionismo naturalista: perspectiva crítica sobre la naturalización de la moral

A las ciencias empíricas corresponde el estudio descriptivo y explicativo de la moralidad humana. A la filosofía moral o ética corresponde la tarea de la fundamentación y aplicación de principios, normas y valores morales. No obstante, tal y como hemos podido ver en los capítulos precedentes, algunas propuestas naturalistas han defendido que es posible una ética naturalizada o una ciencia éticamente normativa. Por ello es necesario hacer una revisión rigurosa de la cuestión. Expondremos y analizaremos los principales errores conducentes a este reduccionismo naturalista. Tras ello cotejaremos si las propuestas mencionadas cometen alguno de esos errores. Nuestra conclusión será que ninguna de ellas resiste la crítica.

Un problema central en la interacción entre ética y ciencias empíricas es el de diferenciar qué competencias corresponden a cada una de ellas en relación con la moral. En cuanto a la ciencia, a esta corresponde la tarea de explicar y describir el sustrato biológico de la moralidad. Se han producido significativos avances desde la biología evolutiva (Darwin, 1979; Huxley, 1894; Ayala, 2006; Ayala y Cela Conde, 2001; Tomasello, 2013; Waal, 1997, 2007; Joyce, 2006, 2014), la sociobiología (Hamilton, 1964; Trivers, 1971, 1985; Wilson, 1975, 1984, 1998), la psicología (Piaget, 1987; Kohlberg, 1992; Hauser, 2008; Tomasello, 2013; Greene, 2003) y, más recientemente, la neurociencia, concretamente desde el campo de la neuroética (Damasio, 2005, 2008; Greene, 2003; Churchland, 2013). Sin embargo, algunas propuestas han pretendido extraer conclusiones éticas de carácter normativo a partir de datos empíricos (Spencer, 1864, 1883; Richards, 1986; Arnhart, 1998; Mora, 2007; Gazzaniga, 2006; Greene, 2003, 2008, 2013, 2014; Churchland, 2008, 2013; Casebeer, 2003; sobre la ciencia como método aplicado a la ética ver: Dewey, 1902, 1945; Ruse y Wilson, 1986). Desde la filosofía moral se ha denunciado que tales posturas incurren en un reduccionismo naturalista, argumentando que si bien la ciencia puede dar cuenta de la base biológica de la moralidad no puede, sin embargo, ofrecer una fundamentación ética de principios, valores y normas morales, esta tarea correspondería a la filosofía moral (Hume, 1988, 1953; Moore, 1983; Gould, 1999; Korsgaard, 2000; Habermas, 2006; Cortina, 2007, 2012a; Salles, 2013, 2016, 2018; Evers, 2010; Ortega Esquembre, 2016; Pallarés-

Domínguez, 2013; Ortiz de Landázuri, 2016; Nagel, 1985; Cartwright, 2010; Teehan y diCarlo, 2004). En la crítica que el discurso filosófico dirige contra estas tentativas naturalistas se han denunciado diversos errores, teniendo como referentes a Hume y a la célebre falacia naturalista de Moore. Dada la vigente actualidad del problema, se hace preciso un análisis riguroso de la cuestión. Nuestros objetivos serán: a) analizar y exponer los diversos errores argumentativos que pueden conducir a un reduccionismo naturalista en materia de ética; b) analizar los argumentos esgrimidos desde el sector naturalista, en defensa de una ética naturalizada o una ciencia éticamente normativa; c) cotejar si estos argumentos y estas propuestas naturalistas cometen los errores previamente expuestos. A la luz de todo ello concluiremos que las propuestas naturalistas de este tipo no resisten la crítica.

Falacias y reduccionismo naturalista

Es preciso realizar un análisis riguroso de los errores denunciados por Hume (1988, 1953) y por Moore (1983), evitando que sean confundidos entre sí, un error común en muchos autores (Gould 1999; Teehan y diCarlo, 2004; Cartwright 2010: 430; Casebeer 2003). Si bien las posturas de ambos guardan una cercanía notable, describen errores lógicos bien diferentes (Curry 2006). Además, muchos de los errores lógicos relativos a deducir un *deber ser* de un *es* no quedan englobados bajo ninguna de las falacias que Hume y Moore denuncian, aunque cometen un error reduccionista del mismo tipo. Esto nos obliga a hacer una revisión de todas aquellas falacias y errores lógicos que pueden intervenir al cometer un reduccionismo de corte naturalista en el ámbito de la moral. Ese será nuestro objetivo en el presente capítulo. Para evitar un uso inadecuado de la terminología podemos englobar dichos errores bajo la categoría de *reduccionismo naturalista* o *falacias es-debe*, en lugar de referir a ellos como “falacia naturalista”, dado que esta consiste en un tipo concreto de error que no ha de confundirse con el resto. No obstante, hemos de tener en cuenta que el término *falacia naturalista* es usado de forma amplia para englobar este tipo de errores, a pesar de no responder de forma rigurosa a las condiciones que dicha falacia reúne según la definición de Moore, por lo que no es extraño encontrar en muchos autores dicha terminología.

Falacia de Hume o falacia del salto es-debe.

El primer autor en denunciar de forma fehaciente que las deducciones de deberes morales basadas únicamente en hechos son falaces fue David Hume, en su *Tratado de la naturaleza humana* (1988), en el que aparece este célebre pasaje:

En todo sistema moral de que haya tenido noticia, hasta ahora, he podido siempre observar que el autor sigue durante cierto tiempo el modo de hablar ordinario, estableciendo la existencia de Dios o realizando observaciones sobre quehaceres humanos, y, de pronto, me encuentro con la sorpresa de que, en vez de las cópulas habituales de las proposiciones: *es* y *no es*, no veo ninguna proposición que no esté conectada con un *debe* o un *no debe*. Este cambio es imperceptible, pero resulta, sin embargo, de la mayor importancia. En efecto, en cuanto que este *debe* o *no debe* expresa alguna nueva relación o afirmación, es necesario que esta sea observada y explicada y que al mismo tiempo se dé razón de algo que parece absolutamente inconcebible, a saber: cómo es posible que esta nueva relación se deduzca de otras totalmente diferentes. Pero como los autores no usan por lo común de esta precaución, me atreveré a recomendarla a los lectores: estoy seguro de que una pequeña reflexión sobre esto subvertiría todos los sistemas corrientes de moralidad, haciéndonos ver que la distinción entre vicio y virtud, ni está basada meramente en relaciones de objetos, ni es percibida por la razón. (Hume 1988: 632-634)

Algunos lectores de Hume han cometido el error de considerar que cualquier deducción de un *es* o un *no es* a un *debe* o un *no debe* moral es, según el autor, falaz (Casebeer 2003). En realidad, Hume está diciendo que no basta *únicamente* con la descripción de hechos para extraer conclusiones sobre lo que *debe* o *no debe ser* (Curry 2006). En el anterior pasaje afirma que “este debe o no debe expresa alguna nueva relación o afirmación”, por lo que se precisan dos cosas al mismo tiempo: a) es necesario que esta nueva relación o afirmación sea observada y explicada, es decir, hace falta la inclusión de una nueva premisa, con contenido ético, que establezca el vínculo entre los hechos y las conclusiones normativas que de ellos se pretenden extraer, y b) que se de razón de cómo es posible que esta nueva relación se deduzca de otras totalmente diferentes, es decir, explicar cómo puede haber una conexión -en la premisa ética- entre el ámbito descriptivo y el ámbito normativo.

Algunos comentaristas de Hume han hecho una lectura más detallada del pasaje anteriormente citado (Quintelier et al., 2011: 34; Sloan Wilson et al., 2003: 670-671, 680; Walter, 2006: 35; Cartwright, 2010: 430; Sober, 1994; Thornhill y Palmer's, 2000; Blackburn, 1998, 2001; MacIntyre, 1959; Mackie, 1980), subrayando que Hume no niega

la posibilidad de deducir un deber moral a partir de hechos, sino de hacerlo *exclusiva o únicamente* a partir de estos, se necesita una segunda premisa que introduzca consideraciones éticas (a).

De este modo, según Hume, un argumento con la siguiente forma sería falaz:

P1. Los seres humanos tienen una predisposición biológica hacia el altruismo.

Conclusión. Es moralmente bueno que los seres humanos actúen altruistamente.

Para que este argumento sea lógicamente válido habría que introducir una segunda premisa que permita que la conclusión se derive de los hechos. Con la siguiente forma el argumento dejaría de ser falaz:

P1. Los seres humanos tienen una predisposición biológica hacia el altruismo.
(Premisa factual)

P2. Es moralmente bueno actuar en consonancia a las predisposiciones biológicas.
(Premisa ética)

Conclusión. Es moralmente bueno que los seres humanos actúen altruistamente.

Que este argumento sea válido en su estructura lógica no implica que sea cierto lo que plantea, para ello se precisa demostrar que las premisas son verdaderas, lo cual garantizaría que también lo es la conclusión. La Premisa 2, como expondremos más adelante, no es verdadera. P2 precisa, además, la otra condición que Hume exige (b), consistente en dar cuenta de cómo la premisa ética puede establecer una conexión entre una cuestión de hecho y un deber moral.

Lo que nos interesa del caso es ofrecer una lectura rigurosa de Hume, pudiendo determinar exactamente qué tipo de error denuncia este. Puesto que este error difiere notablemente de otros, como la falacia naturalista de Moore, llamaré a esta la *Falacia de Hume* (referida así por Curry, 2006: 236) o *Falacia del salto es-debe* (haciendo referencia a cómo el problema ha sido referido tradicionalmente, siendo así, además, como Hume refiere a él: *is-ought gap problem*). A continuación ofrezco una breve descripción de la misma.

Falacia de Hume o falacia del salto es-debe: un argumento será falaz cuando se consideren una o varias premisas factuales como necesarias y suficientes para extraer conclusiones normativas. Para evitar la falacia se precisa a) la introducción de una premisa con contenido ético y b) la justificación de cómo esta premisa puede establecer una conexión entre cuestiones de hecho y deberes morales.

La conexión que exige la condición *b* constituye uno de los objetivos del discurso de la filosofía moral, en la medida en que este aborda el problema de la justificación y la fundamentación ética. Al contrario de lo que muchos han afirmado (Quintelier et al., 2011; Kurtz y Koepsell, 2007; Rosenberg, 2000; Binmore, 2005; Flanagan, 1996; Casebeer, 2003; O. Harman, 2012: 560; Walter, 2006: 39; Pigliucci, 1987; Sober, 1994), la filosofía moral no niega la relevancia de ciertas cuestiones de hecho para la reflexión ética, sino —como Hume— la suficiencia de estas cuestiones para resolver el problema de la fundamentación. Así pues, Hume denuncia no solo la ausencia de una premisa que introduzca un juicio, norma o valor moral (a), sino que también denuncia la ausencia de justificación y fundamentación de los contenidos morales que tal premisa suscribe (b). Pretender que exclusivamente los hechos sean suficientes para extraer conclusiones normativas equivaldría a defender que los hechos son suficientes para fundamentar contenidos morales. Esto es lo que Hume no estaba dispuesto a admitir. Cualquier argumento que pretenda tal cosa es falaz.

Falacia de Moore o falacia definicionista.

Moore en sus *Principia Ethica* (1983) planteó que algunos argumentos que pretenden deducir deberes morales a partir de hechos incurren en una falacia que él acuñó como *falacia naturalista*. Para comprender de forma rigurosa esta falacia hemos de presentar brevemente la propuesta de Moore en los *Principia ethica*. Para este el objeto de la ética es “lo bueno”, más concretamente, lo moralmente bueno. Por tanto, para que la ética pueda constituirse como una ciencia de corte naturalista “lo bueno” debería ser un objeto natural, al que se pueda acceder mediante la experiencia. Sin embargo, Moore considera que lo bueno no es un objeto de esta clase, sino un objeto simple e indefinible del pensamiento. Lo compara con la percepción del color amarillo, que se da como objeto del pensamiento, y no como propiedad de los objetos naturales (Campbell, 1993; Johnston, 1992).

Nuestra primera conclusión, por lo que toca al tema de la ética, es, pues, que hay un objeto del pensamiento, simple, indefinible e inanalizable, con referencia al cual debe definirse. Es indiferente qué nombre le demos a este objeto único, con tal de que reconozcamos claramente qué es y en qué difiere de otros. (Moore, 1983: 19)

Resulta interesante señalar que si bien Moore refiere a este objeto del pensamiento como lo [moralmente] bueno, no excluye la posibilidad de referir a él con otros términos. Esta consideración permite aplicar la falacia naturalista a los casos de reduccionismo en que no se aborde específicamente lo bueno, sino otros conceptos morales afines, como el de justicia o normatividad, en relación a los cuales la ética también se define. Atendiendo a esta consideración podemos hacer un uso más amplio de su terminología.

Atendamos, ahora, al fragmento de los *Principia Ethica* en el cual acuña el término “falacia naturalista”. Es el siguiente:

Puede ser verdad que todas las cosas que son buenas son *también* algo más, tal como es verdad que todas las cosas amarillas producen una cierta clase de vibración lumínica. Y es un hecho que la ética pretende descubrir cuáles son aquellas otras propiedades que pertenecen a las cosas buenas. Pero un enorme número de filósofos han pensado que, cuando nombran esas otras propiedades, están definiendo ‘bueno’ realmente, y que no son, de hecho, ‘otras’ sino absoluta y enteramente iguales a la bondad. A esta postura propongo que se la llame ‘falacia naturalista’, y me esforzaré ahora en darle su justo lugar. (Moore, 1983: 9)

Moore plantea que lo bueno es percibido como objeto del pensamiento, siendo además percibido en relación a un objeto, propiedad o hecho natural. Sin embargo, apunta, muchos autores consideran que ese objeto natural no es algo de lo que se predique lo bueno, bien como propiedad, o bien como causa de la percepción de lo bueno, sino que existe una identidad, en su sentido más riguroso, entre el objeto natural y lo bueno. En tanto que indefinible, cualquier tentativa de definir lo bueno en términos naturalistas resulta falaz (Moore, 1983; Teehan y diCarlo, 2004: 34-35; Hill, 1976).

El error descansaría en considerar que un enunciado como “el placer es bueno” significa “el placer es idéntico a lo bueno” ($P=B$), en lugar de significar “del objeto placer puede predicarse que es bueno” (P_b). También se cometería este error si se identifica una propiedad de un objeto con lo bueno, por ejemplo “el objeto (O) tiene la propiedad de ser placentero (p), y lo placentero es idéntico a lo bueno ($p=B$). En ambos casos, bien sea el

objeto o bien sea una de sus propiedades, el término *es* no predica de estos elementos que sean buenos, sino que lo identifican con lo bueno, como una misma cosa.

Si la falacia solo fuese cometida desde el ámbito naturalista de las ciencias empíricas el término con el cual Moore la acuñó sería adecuado. Pero Moore, además, consideró que también desde la metafísica se comete esta falacia al definir lo bueno identificándolo con objetos de una realidad metafísica o suprasensible.

Las teorías que me propongo examinar pueden dividirse convenientemente en dos grupos. La falacia naturalista implica que siempre que pensemos ‘esto es bueno’, estamos pensando en que la cosa de la que se trata mantiene una relación definida con alguna otra. Pero esta cosa una, con referencia a la cual se define bueno, puede ser o lo que podríamos llamar un objeto natural —algo cuya existencia es admitidamente objeto de la experiencia— o puede ser un objeto del que solo se infiere que ha de existir en un mundo real suprasensible. Me propongo tratar por separado estos dos tipos de teorías éticas. Las del segundo tipo podrían convenientemente llamarse ‘metafísicas’, y pospondré su consideración hasta el capítulo IV. [...] Debe observarse que la falacia con referencia a la cual defino la ‘ética metafísica’ es de la misma clase, y no le daré más nombre que el de falacia naturalista. (Moore, 1983: 36-37)

Si la misma falacia es cometida tanto en una aproximación empírica y naturalista a la moral como desde una aproximación metafísica, está claro que lo distintivo y característico de esta falacia no es su tentativa naturalista, sino alguna otra cosa. Coincido con algunos comentaristas de Moore (Frankena; Warnock; Teehan y diCarlo, 2004: 35) en que un nombre más adecuado podría ser el de “falacia definicionista” —en el original *definist fallacy* (Frankena)—, ya que lo característico tanto desde el naturalismo como desde la metafísica sería definir lo bueno como idéntico a un objeto o propiedad natural o a un objeto o propiedad metafísico de una realidad suprasensible, respectivamente. Propongo, pues, seguir esta sugerencia, considerando terminológicamente más adecuado referir a esta falacia como *Falacia de Moore* o *Falacia definicionista*. Esta podría darse en dos modos, el de *falacia naturalista* —siendo este término más restrictivo que el uso que Moore hace de él—, si la definición se produce en relación a un objeto o propiedad natural, y el de *falacia metafísica*, si la definición se produce en relación a un objeto o propiedad metafísica de una realidad suprasensible. Una definición sumaria de estas falacias sería la siguiente:

Falacia de Moore o *falacia definicionista*: un argumento será falaz cuando pretenda extraer conclusiones normativas a partir de hechos en virtud de definir

lo bueno como idéntico a un objeto, hecho o propiedad. Puede darse en dos modos: a) *falacia naturalista*, cuando el objeto en relación al cual se establece la definición es un objeto natural; b) *falacia metafísica*, cuando el objeto en relación al cual se establece la definición es un objeto metafísico propio de una realidad suprasensible.

En los casos en que se comete la falacia definicionista naturalista, dado que el objeto o propiedad identificado con lo bueno sería natural, su descubrimiento, descripción y explicación serían competencia de las ciencias empíricas, y, por consiguiente, también la fundamentación ética, que resultaría ser, por estos motivos, una cuestión de hecho. Sin embargo, una vez desvelada la naturaleza falaz de este tipo de argumentos, toda postura que se asiente sobre ellos queda invalidada.

Falacia de ambigüedad por equívoco.

Las falacias de ambigüedad son aquellas que se producen como fruto de emplear uno o varios términos con diferentes significados o sentidos. Entre las falacias de ambigüedad encontramos las siguientes: falacia por equívoco, falacia por anfibología, falacia de énfasis y falacia por división. De todas ellas atenderemos a la primera, ya que puede originar un reduccionismo naturalista en el ámbito de la ética.

La falacia de ambigüedad por equívoco, también llamada falacia por homonimia, tiene lugar cuando se hace un uso indiferenciado de dos términos homónimos (diferentes términos con igual forma y distinto significado) o de varias acepciones de un término polisémico (un mismo término con varios significados), generando un error en la argumentación (García Damboreana, 2000). Veamos un ejemplo:

Consideremos algunos de los significados del término “fin”. En su sentido teleológico puede designar el perfeccionamiento de algo. Otro de sus significados es el de acabamiento. Usar equívocamente el término, en oposición a usarlo unívocamente, puede generar argumentos falaces como el siguiente:

P1. El fin de una cosa es su perfección. (sentido teleológico)

P2. La muerte es el fin de la vida. (sentido de acabamiento)

Conclusión. La muerte es la perfección de la vida.

Esta falacia puede conducir a posturas reduccionistas en el ámbito de la moral. Algunos de los términos más comunes que originan este error son: a) *valor*, hay que distinguir lo que es valioso en términos biológicos, por ejemplo en la selección natural, en la preservación de la vida, en su buen desarrollo, etc., de los valores morales, que no coinciden necesariamente con esos otros valores biológicos; b) *bueno*, del mismo modo, lo que puede ser bueno en términos biológicos difiere en significado de lo que puede ser bueno en el ámbito moral; c) *deber*, en un sentido prudencial puede usarse este término, por ejemplo, para indicar el medio necesario para la consecución de un fin (“si quieres curar tu dolor, *debes* ir al médico”), o para señalar alguna predisposición natural o cultural (“la naturaleza dicta que la madre *debe* cuidar de las crías”), este uso del término difiere del uso que se hace en el ámbito moral, el deber moral no tiene por qué coincidir con las pautas de conducta que puedan ser descritas o prescritas en ámbitos no morales, aunque también se use el termino *deber* para referirlas.

Una definición sumaria de esta falacia sería la siguiente:

Falacia de ambigüedad por equívoco o *falacia por homonimia*: un argumento será falaz cuando la conclusión se derive de las premisas por la ambigüedad derivada de hacer un uso sinonímico de dos términos homónimos o de diferentes acepciones de un término polisémico, en función de lo cual hay coincidencia en la forma pero no en el significado.

Falacia por transposición categorial.

La falacia de ambigüedad por equívoco —en la medida en que se produce por el uso de diferentes términos homónimos o diferentes acepciones de un término polisémico, pudiendo estos pertenecer a diferentes ámbitos o esferas, como el de la naturaleza y el de la filosofía moral— puede ser causa de que otro error lógico tenga lugar, el de la falacia por transposición categorial. Esta se produce como fruto de aplicar una categoría de un ámbito o esfera a otro que es totalmente distinto. Veamos un ejemplo de cómo un error de este tipo puede generar un reduccionismo naturalista en el ámbito de la moral.

P1. La evolución ha inscrito ciertos valores en nuestra psicobiología. (Valor psicobiológico)

P2. Los valores son criterios para guiar moralmente nuestra conducta. (Valor moral)

Conclusión. Nuestra psicobiología ofrece criterios para guiar moralmente nuestra conducta.

Lo biológicamente valioso y lo moralmente valioso no refieren al mismo tipo de *valor*, el argumento arriba expuesto descansa sobre una falacia de ambigüedad por equívoco que, a su vez, genera una transposición categorial, ya que se extrapolan consideraciones del ámbito biológico y natural al ámbito filosófico de la fundamentación ética.

Podemos definir de forma sumaria este error del siguiente modo:

Falacia por transposición categorial: un argumento es falaz si sus premisas aplican una categoría o concepto propio de una esfera o ámbito a otros que son diferentes. Puede producirse como resultado de una falacia de ambigüedad por equívoco.

Falacias interconectadas.

La relación existente entre estos errores provoca, a menudo, que varios de ellos se cometan simultáneamente. Describir estas relaciones es preciso para obtener una mejor comprensión de por qué algunos argumentos reduccionistas resultan falaces.

Una de estas relaciones ya ha sido descrita, la existente entre la falacia de ambigüedad por equívoco y la falacia por transposición categorial. En la medida en que dos términos homónimos o dos acepciones de un término polisémico sean usados indistintamente, además de producirse una falacia de ambigüedad por equívoco, puede darse un salto de una esfera o ámbito a otro totalmente distinto, produciendo así, además, una falacia por transposición categorial.

La falacia de Moore también está relacionada con estas dos. La confusión entre un *es* de identidad y un *es* de predicación puede dar lugar a cometer la falacia de Moore. Al hacer un uso indistinto de dos acepciones diferentes del término *es* se está cometiendo, además de una falacia definicionista o falacia de Moore, una falacia de ambigüedad por equívoco, que a su vez provoca una falacia por transposición categorial. No solo se

producen estos errores en relación al término *es*, la confusión de los términos *valor* y *bueno* en el ámbito biológico respecto a sus homónimos en el ámbito de la moral puede conducir también a una falacia definicionista naturalista basada en una falacia de ambigüedad por equívoco, dando como resultado una falacia por transposición categorial.

Todas estas falacias se hallan, además, conectadas con la falacia de Hume. Este especifica que no se pueden deducir deberes morales única y exclusivamente a partir de hechos, es necesaria (a) la introducción de una premisa ética y (b) la justificación de cómo el contenido de esta premisa puede conectar los hechos con los deberes morales. Los argumentos basados en cualquiera o cualesquiera de las falacias a las que hemos atendido pretenden a menudo cumplir la segunda condición que Hume exige (b), justificando la conexión entre los hechos y el deber moral. De este modo no habría falacia de Hume, dejaría de darse, puesto que no se trataría de deducir el deber moral exclusivamente de hechos, se habría introducido una premisa ética que incorporaría la justificación exigida. El problema viene dado en que esta premisa descansará en un argumento falaz por introducir en ella una falacia de Hume, una falacia de Moore, una falacia de ambigüedad por equívoco, una falacia por transposición categorial, o una combinación de algunas de las anteriores.

El hecho de que todas estas falacias puedan estar conectadas en algunos argumentos reduccionistas dificulta el análisis de estos. Por ello para hacer una crítica rigurosa de dichos argumentos resulta necesaria la tarea que hemos llevado a cabo, consistente en clarificar y explicar las principales falacias del espectro *es-debe*, así como la posible relación entre ellas

La naturalización de la ética sometida a crítica

Existe un sector de investigadores que defiende la posibilidad de una ética naturalizada o de una ciencia éticamente normativa. En defensa de su postura se han manejado diversos argumentos, a continuación analizaremos algunos de los más significativos, dividiendo estos en dos tipos: a) cuestionamiento de las críticas recibidas, poniendo en entredicho la validez y corrección de las mismas; y b) propuestas de una ética naturalizada o de una ciencia éticamente normativa, que o bien refutan las críticas contra el reduccionismo, o bien las aceptan pero pretenden suponer una excepción a las

mismas. En las dos siguientes secciones mostraré que ninguno de estos argumentos resiste la crítica.

Argumentos que cuestionan las acusaciones de reduccionismo.

La ciencia informa a la ética. Una de las objeciones contra los críticos del reduccionismo es que el naturalismo metodológico no pretende resolver el problema de la fundamentación, sería apropiado para abordar la investigación científica en el ámbito de la moralidad, pudiendo resultar normativo en la medida en que la ciencia pueda informar a la ética (Quintelier et al., 2011; Kurtz y Koepsell, 2007; Rosenberg, 2000; Binmore, 2005; Pigliucci, 1987; Sober, 1994), tanto en cada problema concreto, como en relación a las posibilidades y límites que nuestra naturaleza impone a nivel conductual (Flanagan, 1996; Casebeer, 2003; Harman, 2012: 560). Desde esta postura se ha apelado a la llamada “falacia anti-naturalista”, según la cual se ha de reconocer que, aunque todos los hechos naturales no son relevantes para el discurso moral, todos los hechos que son relevantes para tal discurso son naturales (Casebeer, 2003; Walter, 2006). En respuesta a ello cabe señalar que la falacia anti-naturalista no es cometida por la crítica contra el reduccionismo naturalista, ya que en realidad tal crítica reconoce que todos los hechos que son relevantes para el discurso moral son hechos naturales, sin embargo subraya que estos hechos no son suficientes para fundamentar la ética. Moore no niega el papel informativo de los hechos para la ética, y tampoco la posibilidad de una ciencia sobre lo moral, como algunos autores sostienen (Walter, 2006: 39). Tampoco Hume lo niega, al exigir una premisa que justifique la conexión entre hechos y deberes morales no niega la importancia de los primeros, es más, reconoce su papel informativo. Desde la crítica al reduccionismo naturalista, por tanto, se reconoce que los hechos resultan relevantes para la ética, pero esto no significa que sean suficientes (Teehan y diCarlo, 2004).

La ciencia puede resultar éticamente normativa, por reducción al absurdo. Muchos filósofos consideran que el naturalismo no es suficiente para ofrecer criterios moralmente normativos, mientras que la filosofía sí, sin embargo, según sostienen algunos naturalistas, la filosofía no ha podido resolver el problema de la fundamentación (Curry, 2006: 240-242). Con ello se pretende adjudicar a la ciencia tal tarea. Este sería un argumento por reducción al absurdo, con la siguiente forma: supongamos que la filosofía es capaz de resolver el problema de la fundamentación, no lo ha resuelto, luego no corresponde esta tarea a sus competencias. Hay varias cosas que objetar. En primer lugar,

hay que señalar la existencia de diversas teorías éticas que proponen criterios normativos, por tanto el problema puede haber sido resuelto por estas, aunque no haya unanimidad al respecto. Además, suponiendo que la filosofía no lo haya resuelto, eso no significaría que no sea capaz de resolverlo. Por último, hay que añadir que si se demostrase tal incapacidad —cosa que no se ha hecho—, a lo sumo podría concluirse que no corresponde a la filosofía resolver este problema, lo cual no implicaría que corresponda a la ciencia.

Argumento sobre la naturaleza humana en Hume. Una de las objeciones que se han hecho respecto a Hume es que su famoso pasaje no ofrece en realidad un argumento contra el reduccionismo naturalista, sino una aportación a su teoría sobre la naturaleza moral humana (Ruse; Joyce, 2000; Cartwright, 2010: 430; Walter, 2006: 36; Curry, 2005, 2006). Hume propone que el vicio y la virtud no se hallan en los objetos o hechos naturales, sino en la percepción que el sujeto tiene de estos. El sentimiento, además, resulta necesario para motivar la conducta moral, la razón no es causa suficiente para ello (Hume, 1988: 632-633, 636; Siurana, 2011b: 139-160). Una inadecuada lectura de Hume considera que la deducción de deberes morales a partir de hechos estaría injustificada dado que estos no son por sí mismos virtuosos o viciosos. La justificación del deber moral descansaría, según esta interpretación, en las motivaciones psicológicas del ser humano, que no escapan, en realidad, al ámbito de la naturaleza. Sin embargo, Hume insiste en aclarar esta cuestión repetidas veces en otros pasajes de su *Tratado de la naturaleza humana*, especificando que la moralidad, en su sentido más fundamental, no consiste en una cuestión de hecho (Hume, 1988: 626, 632), la justicia, dice, descansa en instancias artificiales, aunque necesarias, como la educación o las convenciones humanas (Hume, 1988: 651, 652, 657, 658, 667, 704).

Lo bueno como metafísico en Moore. Una de las críticas esgrimidas contra Moore es que, al negar que lo bueno corresponda a una propiedad o a un objeto natural, este ha de recurrir a entidades metafísicas (Quintelier et al., 2011: 35; Putnam, 1981; Teehan y diCarlo, 2004; Walter, 2006: 34; Warnock, 1978). Esta interpretación de Moore no es correcta, este sugiere que lo bueno es percibido como un objeto del pensamiento que, al ser indefinible y simple, se produce por intuición (para una crítica detallada al concepto de lo bueno como indefinible en Moore ver: Joyce, 2000; Frankena, 1973; G. Harman, 1977: 19; Putnam, 1981). Para Moore lo contrario de *natural* es *intuitivo*, no *sobrenatural* (Curry, 2006: 239). Que el pensamiento sea un objeto natural no significa que el objeto

del pensamiento sea o refiera a un objeto o propiedad natural (Curry, 2006; Radcliffe Richards, 2001: 246). Asimismo, no hace falta apelar a una realidad metafísica sobrenatural para que la mente pueda percibir objetos, propiedades o conceptos que no tienen una correspondencia con elementos del mundo externo, especialmente si tratamos con conceptos que no describen el mundo, sino que ofrecen pautas normativas acerca de cómo comportarnos en él. De esta mala lectura muchos críticos concluyen, erróneamente, que Moore y sus simpatizantes denuncian cualquier sistema de ética naturalista, incluyendo las propuestas puramente descriptivas que se han hecho desde la ciencia, especialmente las de corte evolucionista (Quintelier et al., 2011: 36; Richards, 2000; Doris y Stich, 2005). En contra de esta afirmación hay que aclarar que las críticas contra el reduccionismo naturalista no pretenden desacreditar este tipo de aportaciones científicas, meramente descriptivas y explicativas, tal como han advertido muchos autores (Ruse, 2008; Binmore, 2005; Farber, 1994; Woolcock, 1999; Rosenberg, 2000; Gaulin and McBurney, 2001; Barrett et al., 2002; Sloan Wilson et al., 2003; Gould 1999).

La verdad analítica de lo bueno en Moore. En la medida en que Moore considera que lo bueno es percibido como un objeto del pensamiento, algunas críticas, desde una perspectiva lógica, han alegado que no se puede obtener una verdad analítica en relación a ese concepto de lo bueno, salvo que se plantee tautológicamente, en cuyo caso no aportaría conocimiento alguno al respecto (Casebeer, 2003; Flanagan, 2008; Holmes, 1990; Wilson, 1975). El error en esta crítica descansa en la creencia de que Moore y otros críticos pretenden establecer una verdad moral en términos analíticos. De hecho, la ética, a excepción de algunas propuestas realistas, no busca encontrar enunciados morales verdaderos, sino justificables (Korsgaard, 2000; Apel, 1985; Habermas, 2006; Cortina, 2007), la búsqueda de la verdad mediante un método analítico no supone uno de sus objetivos. El sentido en el que Moore habla de verdad puede, por supuesto, ser sometido a crítica (Quinteliert et al., 2011), pero al margen de eso la corriente crítica sobre el reduccionismo naturalista no plantea que la ética busque una verdad moral (para una crítica más detallada a Moore en esta cuestión ver: Walter, 2006: 41).³²

³² Más comentarios sobre Moore en: Regan, 1986; Farber, 1994; Woolcock, 1999; Rosenberg, 2000; Kitcher, 1994; Richards, 1986; Arnhart, 1998; Held, 2002; Flew, 1967, 1978; Sturgeon, 2003; Walsh, 2008.

Argumentos a favor de una ética naturalizada o una ciencia éticamente normativa.

Ética evolucionista. Uno de los autores clásicos en esta materia, foco de crítica del propio Moore (1983), es Herbert Spencer (1964, 1883), quién trató de extraer conclusiones normativas atendiendo a la propuesta evolutiva de Darwin (comentado en: Delaney, 2009; Cartwright, 2010; Quintelier et. al., 2011). Su propuesta de una ética de corte evolucionista alentó modelos políticos de darwinismo social (Morgans; Haeckel; Bernstein; Babel's y Virchow, citados en: Richards, 1986: 265-268). Spencer plantea la evolución como un progreso, cuyos resultados tienden a una superioridad y perfeccionamiento, en función de lo cual las pautas biológicas de conducta serían un criterio normativo para el comportamiento moral. Spencer comete varios errores. Su postura descansa en una mala comprensión de qué es la evolución, esta no tiene una dirección, ni produce seres cada vez mejores o superiores en sentido moral, los únicos criterios para establecer una mejora o superioridad serían en relación al éxito de supervivencia y reproducción en adecuación al medio (Dennett, 1995). Spencer no comete la falacia de Hume, ya que introduce una premisa ética, estableciendo una conexión entre la evolución y lo normativo. Sin embargo, en esta premisa comete una falacia de Moore, denunciada por este en sus *Principia Ethica*, al definir lo bueno en términos evolutivos. Una versión similar de este error, esta vez cometido desde el ámbito de las neurociencias, es llevada a cabo por Mora (2007: 72, 159), Gazzaniga (2006: 17-18, 179) y Greene (2013), quienes consideran que las pautas de conducta que la evolución ha inscrito en nuestra biología permiten fundamentar unas normas morales de forma universal, puesto que toda la especie comparte el órgano que las contiene, el cerebro.

Valores biológicos como valores morales. Muchos autores consideran que la biología predispone ciertos valores que resultarían moralmente normativos (Casebeer, 2003: 34; Flanagan; Cartwright, 2000: 328; McShea y McShea, 1990, 1999; Churchland, 2013; Curry, 2006: 235, 237-238). Churchland, además, reconoce que estos valores serían maleables durante la epigénesis y la ontogénesis. Curry sostiene, por su parte, que una psicología actualizada en la investigación de lo moral implica una metaética actualizada, siendo la filosofía moral una rama de la biología, la psicología y la antropología en la medida en que los valores pueden definirse en términos adaptativos relativos a la cooperación. La confusión entre valores biológicos y valores morales comete, en todos

estos casos, una falacia de ambigüedad por equívoco, que conduce a una transposición categorial, por confundir dos tipos de *valor* distintos, el biológico y el moral, extrapolando el papel conductual de los primeros al ámbito normativo de la ética.

Ser moral es una cuestión de hecho. Richards ofrece una “Versión Revisada” de las propuestas científicas en materia de ética, atendiendo especialmente a Darwin y Wilson, con el objetivo de ofrecer criterios de justificación moral basados en las aportaciones de la biología evolutiva, sin cometer ninguna falacia por ello (Richards, 1986: 272; comentado por: Joyce, 2000; Cartwright, 2010). Su propuesta plantea que “la evolución, como cuestión de hecho, ha construido a los seres humanos para actuar por el bien de la comunidad”, en ello consistiría “ser moral”. Así que, puesto que “los seres humanos son seres morales —una condición inevitable producida por la evolución— deben actuar por el bien de la comunidad” (Richards, 1986: 289). Richards comete varios errores. Uno de ellos es la falacia de Moore, al identificar lo moralmente bueno con cierto tipo de conductas naturales, concretamente las altruistas y las que persiguen el bien común (comentado en: Hughes, 1986). El criterio que usa para definir lo bueno o lo moralmente correcto también se basa en una falacia, en concreto de ambigüedad por equívoco. Considera que *ser moral* es algo que la biología evolutiva puede describir, sin embargo *ser moral* en sentido biológico, tener la capacidad para la moralidad, no es lo mismo que *ser moral* en tanto que actuar éticamente. Confundir estos dos modos de *ser moral* lleva a Richards a esta falacia por equívoco, que además conduce a una falacia por transposición categorial, puesto que traslada el esquema natural de la conducta moral humana al esquema normativo de la ética, que es bien diferente. Por tanto, queda sin justificar desde su propuesta naturalista por qué las predisposiciones biológicas hacia el altruismo y la búsqueda del bien común son éticamente deseables. Respecto al término *deber* comete los mismos errores, considerando que la constrictión biológica a actuar de cierto modo es sinónima a la constrictión normativa impuesta por la ética. No comete la falacia de Hume ya que introduce una premisa ética, pretendiendo justificar la conexión entre hechos y deberes morales, pero esta premisa resulta ser falaz, por los motivos que hemos apuntado.

Elementos naturales intrínsecamente normativos en sentido moral. Han sido muchos los autores que han considerado que existen elementos naturales éticamente normativos, tales como los deseos (Arnhart, 1998: 17; comentado por: Lemos 2000), el

placer (Singer 1981, 2001; para un tratamiento más detallado de la cuestión ver clásicos del utilitarismo: Bentham, 2008; Mill, 1968) o el florecimiento (Casebeer, 2003; MacIntyre, 2001). Estas posturas han sido ampliamente debatidas (Korsgaard, 2000; Moore, 1983). En ellas se producirá una falacia de Hume si se apela únicamente a los hechos como fuente de la normatividad. Sin embargo, es frecuente introducir una premisa ética que pretenda establecer una conexión entre hechos y deberes morales, la más común consiste en definir el deseo, el placer o el florecimiento como lo bueno, o como criterio suficiente para la fundamentación ética, incurriendo así en una falacia de Moore. De hecho, algunos aspectos del utilitarismo y el hedonismo son foco de esta crítica por parte del propio Moore en sus *Principia Ethica*.

Los juicios morales correctos son una cuestión de hecho. Arnhart sostiene que “juicios morales correctos son juicios factuales sobre los patrones típicos de los sentimientos morales de la especie en circunstancias específicas” (Arnhart, 2005: 208; comentado en: Cartwright, 2010; Lemos, 2000). Según esto, ciertos patrones conductuales de la especie en determinados contextos pueden ofrecer criterios normativos, produciendo juicios morales correctos. Si bien el funcionamiento del juicio moral puede describirse como predisposición conductual en determinados contextos, ello no ofrece criterios para justificar por qué la solución ofrecida por el juicio es la moralmente correcta. Arnhart no comete la falacia de Hume, en la medida en que introduce una premisa ética que pretende conectar la esfera de los hechos con la del deber moral. Su argumento resulta falaz precisamente en los términos en que esta premisa introduce tal conexión. El error que comete consiste en una falacia de Moore, ya que considera que lo bueno puede describirse como el juicio moral producido por ciertas predisposiciones biológicas en determinados contextos. Un argumento similar, aunque con algunos matices que lo diferencian, es propuesto por otros autores como Churchland (2013) —a la cual atenderemos en el siguiente párrafo—, y, comentando la obra de esta, O. Harman (2012).

Inferir deberes de hechos como solución funcional. Refiriendo a Hume, Churchland plantea que en un sentido más amplio de *inferir* que el de derivación lógica, sí es posible deducir conclusiones sobre cómo debemos actuar a partir de hechos (Churchland, 2013: 11-21). Expone que la resolución de problemas concretos en contextos determinados se halla sujeta a una “satisfacción de restricción”, ya que estos

problemas vienen sujetos a una serie de hechos, entre ellos nuestra biología y el contexto en el que surge un problema. Por ello, argumenta, quizás la solución no dependa de una argumentación lógica, sino de ofrecer una respuesta adecuada y factible al problema, aunque esta respuesta no sea la mejor solución en sentido moral es, al menos, una solución que funciona. Hay varios problemas en el planteamiento de Churchland. En primer lugar, las consideraciones que hace sobre Hume no son del todo adecuadas, precisamente el carácter lógico de la crítica humeana es el que inhabilita las posturas que pretenden deducir deberes morales únicamente a partir de hechos. Considerar que una acepción más amplia de *inferencia* permite evitar el reduccionismo que denuncia Hume es adular la crítica de este, la cuestión a la que nos enfrentamos no consiste en ofrecer soluciones funcionales a un problema, sino en analizar la posibilidad de argumentar lógicamente a favor de una deducción de deberes morales exclusivamente a partir de hechos. Además, Churchland está eludiendo el núcleo del problema moral, los hechos pueden informar sobre cómo *debemos* actuar en un sentido prudencial e instrumental (ir al dentista ante un dolor de muelas), pero este no es un *deber* en sentido moral. Se está produciendo aquí una falacia de ambigüedad por equívoco. Este mal uso del término *deber* es cometido también por Richards (1986), tal y como denuncia Joyce (2000) (sobre diferentes usos del término deber consultar: G. Harman, 1975; Mackie, 1977: 74). Además, Churchland considera que la naturaleza ofrece ciertos fines como moralmente deseables, incurriendo así en una falacia de Moore. Su propuesta trata de argumentar que existe una conexión entre hechos y deberes morales, por lo cual no comete una falacia de Hume, pues habría una premisa ética, sin embargo esta premisa es falaz por otros motivos, como acabamos de exponer.

La deontología como proceso psicológico. En sus investigaciones neurocientíficas sobre el juicio moral, Greene introduce algunas consideraciones en torno a la falacia naturalista, para someterla a crítica, considerando que la psicología moral puede resultar normativamente relevante (Greene, 2003, 2008a, 2014a). Según su teoría del proceso dual en el juicio moral existen dos tipos de juicios morales, unos se producen intuitiva y automáticamente, a los que llama *deontológicos*, y otros se producen consciente y controladamente, a los que llama *consecuencialistas*. Con esta terminología está creando términos homónimos a los filosóficos, con el objetivo de ser utilizados en el campo científico de la psicología. El problema es que define la deontología de forma naturalizada y poco rigurosa, como proceso subyacente a los juicios *deontológicos*, tratando de

desacreditar su fuerza normativa como teoría ética (para una crítica más detallada ver: Gracia, 2016, 2017, 2018). Confundir la esfera en la que opera la psicología con la esfera en la que opera la filosofía moral supone un error de transposición categorial, basado a su vez en una falacia de ambigüedad por equívoco.

Tras el análisis y revisión de los principales argumentos de cada postura hemos puesto en evidencia que los defensores de una ciencia éticamente normativa o de una ética naturalizada, o bien se basan en una mala comprensión de las críticas recibidas, o bien cometen algún error lógico, siendo falaces. Por tanto, a la luz de nuestra investigación, ninguna de las propuestas actuales de una ética naturalizada o de una ciencia éticamente normativa resiste la crítica. El problema de la fundamentación y justificación de valores, normas y principios morales, así como el de su aplicación, serían competencia de la filosofía moral o ética, no de la ciencia. Esta, no obstante, sí resulta apropiada para informar al discurso filosófico, así como para ofrecer un informe descriptivo y explicativo de la naturaleza moral humana.

PARTE III: HACIA UNA METODOLOGÍA CRÍTICA E INTERDISCIPLINAR

En este sentido, Richard Rorty declara un desdoblamiento gramatical en nuestros vocabularios explicativos entre aquellos que dirigen su mirada hacia las causas observables y aquellos otros que la dirigen hacia los motivos racionales, resultado, todo ello, de una adaptación funcional de nuestra especie a dos entornos distintos, el natural y el social. La irreductibilidad de un juego de lenguaje al otro no debería intranquilizarnos más que el hecho de que un instrumento no sea sustituible por otro.

Jürgen Habermas, *Entre naturalismo y religión*

En los capítulos anteriores hemos mostrado cómo una naturalización de la moral es posible solo parcialmente. El método de las ciencias empíricas permite abordar adecuadamente el hecho de lo moral, en tanto que la moralidad se asienta sobre estructuras psicobiológicas, cuyo origen, desarrollo y funcionamiento pueden ser descritos y explicados desde el paradigma naturalista. Existe, por tanto, una naturaleza moral en el ser humano, cuyo estudio corresponde a diversas disciplinas científicas, entre ellas la neurociencia, la psicología, la biología, la sociología, la etología, etc. Sin embargo, tal y como hemos señalado repetidas veces, existen casos en los que los científicos exceden sus competencias, pretendiendo naturalizar algunos aspectos de la moralidad que no corresponde abordar a su método, tales como la fundamentación de normas, valores y principios morales. La filosofía resulta necesaria para resolver problemas como el de la fundamentación ética, así como para determinar qué significan conceptos morales como justicia, bueno, obligación y responsabilidad, entre otros.

Por este motivo, en esta parte de la investigación atenderemos a la necesidad de la filosofía moral en el ejercicio de la neuroética. Para ello analizaremos cómo la filosofía propone resolver aquellos problemas relativos a lo moral cuya investigación escapa a las competencias de la ciencia. Entre las diversas teorías éticas centraremos nuestra atención en las propuestas más relevantes de la Teoría Crítica y de la ética dialógica o discursiva, en cuya tradición se inscribe nuestra investigación. Concretamente nuestro estudio es heredero y se suma a las aportaciones de la ya consolidada Escuela de Valencia, caracterizada por retomar las éticas dialógicas en clave cordial (Cortina, 2009). Esta corriente de pensamiento, la dialógica, se caracterizan no solo por ofrecer una propuesta

adecuada en torno al problema de la fundamentación ética, también esclarece las condiciones en las que operan tanto la ética como las ciencias naturales, estableciendo sus condiciones de posibilidad, sus competencias, sus similitudes y sus diferencias. Si bien la reflexión atenderá a diversos autores y obras relevantes, el hilo conductor del estudio en los próximos capítulos vendrá marcado por cuatro figuras que considero clave, cuyo análisis permitirá esclarecer las cuestiones más acuciantes en torno al tema que aquí nos ocupa. En orden, estos autores serán Christine Korsgaard, Karl-Otto Apel, Jürgen Habermas y Adela Cortina. La propuesta de una ética de la razón cordial, de Cortina, reúne las aportaciones de los autores que la preceden, replanteándolas y completándolas con la inclusión de la dimensión cordial de la razón, que se produce siempre encarnada, y no de forma pura, en el discurso moral. Aplicando las aportaciones de estos autores al ámbito de la neuroética propondré, en la última parte de la investigación, las claves para una metodología adecuada para esta disciplina, bajo el rótulo de una neuroética de la razón cordial.

Capítulo 7. El problema de la fundamentación y el origen de lo normativo

La cuestión por los fundamentos últimos de las normas, valores y principios morales es posiblemente la más central para la filosofía moral. Christine Korsgaard sitúa este problema en el centro de su pensamiento y se propone dar una respuesta al mismo. Si bien considero que la autora no logra resolver todas las dificultades concernientes a esta cuestión, hace algunas aportaciones que son muy necesarias para ejercer una reflexión ulterior. Por ello resulta conveniente atender a ella en primer lugar.

Korsgaard, heredera del pensamiento kantiano, lleva a cabo una propuesta dirigida a superar las dificultades derivadas del carácter monológico de este. En sus investigaciones indaga sobre el origen de la obligación moral y de lo normativo. *Las fuentes de la normatividad* (2000) supone uno de sus mayores esfuerzos por hallar el origen de lo normativo para, así, resolver el problema filosófico de la fundamentación ética (Korsgaard, 2000: 14-15).

La pregunta normativa, el ámbito explicativo y el ámbito justificativo de la moral

La primera cuestión que hemos de plantearnos es el modo en que una prescripción moral es normativa. Las normas morales, o reglas éticas —en terminología de la autora—, son normativas en la medida en que imponen una exigencia moral, a la cual deberá ajustarse la acción si queremos obrar éticamente. Son, por tanto, un tipo de enunciado o proposición distintos a los propios de la ciencia, que solo podrían describir y explicar cómo nuestra conducta se regula, sin imponer por ello una exigencia de cómo debe regularse en sentido moral (Buchanan, 2009; Apel, 1985; Habermas, 2006, 2010; Cortina, 2007, 2012a; Aranguren, 1994). Teehan y diCarlo (2004) exponen esta diferencia de competencias entre ciencia y filosofía analizando un texto de Gould (1999), en el cual este refiere como distintos *magisteria* los dominios de la ciencia y de la filosofía moral.

La cuestión que queremos resolver es por qué conceptos morales como virtud, justicia o bien tienen una fuerza normativa, por qué están capacitados para imponer una exigencia moral que guíe y limite la acción. En relación a ello Korsgaard plantea una pregunta, que ella misma acuña como “la pregunta normativa”, la cual indaga en el origen

de la normatividad, preguntando qué justifica las pretensiones normativas de la moralidad.

Cuando buscamos un fundamento filosófico de la moralidad no estamos simplemente tratando de encontrar una explicación de las prácticas morales: también estamos preguntando qué justifica las exigencias de moralidad. Llamo a esto “la pregunta normativa”. (Korsgaard, 2000: 22)

Korsgaard advierte del riesgo de confundir la pregunta normativa con otras cuestiones propias de la filosofía moral. Es preciso aclarar bien en qué consiste esta pregunta, que indaga en los fundamentos últimos del deber y de la obligación moral. Una vez hecho esto será posible analizar cómo distintas propuestas y corrientes filosóficas han tratado de ofrecer una respuesta. Qué metodología ha de adoptar la neuroética para proceder en sus investigaciones dependerá de cómo se formule la pregunta normativa y de qué peso demos a las diferentes corrientes que han tratado de responder a ella. En la medida en que la neuroética es una rama interdisciplinar, será preciso delimitar los ámbitos explicativo y justificativo de la moral que Korsgaard distingue. La distinción entre estas dos dimensiones de lo moral ya ha sido abordada ampliamente en capítulos anteriores y profundizaremos en ella en los próximos capítulos (Buchanan, 2009; Apel, 1985; Habermas, 2010, 2006; Cortina, 2007, 2009, 2012a; García-Marzá, 2016; Aranguren, 1994; Gould, 1999; Hume 1988, 1953; Moore 1983; Salles 2013, 2016, 2018; Evers 2010; Ortega Esquembre 2016; Pallarés-Domínguez 2013; Ortiz de Landázuri 2016; Nagel 1985; Cartwright 2010; Teehan y diCarlo 2004).

Korsgaard considera que es preciso aclarar bien los términos de la pregunta normativa, así como el contexto que la rodea. Para ello, en primer lugar, diferencia dos grandes esferas en lo que concierne a los conceptos éticos que los seres humanos manejamos (bien, deber, obligación, virtud, justicia, etc.). De un lado, dice, encontramos una *teoría de los conceptos morales* y de otro encontramos *los efectos prácticos y psicológicos de las ideas morales*. La teoría de los conceptos morales trata de responder tres cuestiones según la autora, a saber: a) qué significan o qué contienen los conceptos morales, b) a qué se aplican y c) de dónde vienen. En cuanto a los efectos prácticos y psicológicos de las ideas morales estos refieren a la función práctica que los conceptos morales desempeñan en la vida humana, algunos hechos destacables a este respecto son el sentimiento de obligación moral, las motivaciones psicológicas de la acción moral, o la consideración de corrección moral respecto a una acción. En relación a esta esfera

práctica y psicológica hemos de distinguir dos criterios que afectan a la propia teoría de los conceptos morales, uno de ellos es un criterio de *adecuación explicativa*, el otro es un criterio de *adecuación normativa o justificativa*. Estos dos criterios serán clave a la hora de distinguir las competencias metodológicas que son propias de las ciencias empíricas y de la filosofía moral, respectivamente, cuando se ocupan de abordar el estudio de la moralidad, especialmente cuando lo hacen de forma conjunta, como sucede en campos de estudio interdisciplinarios como la neuroética. Así pues, estos criterios dividen la investigación de lo moral en dos ámbitos, uno que concierne a la explicación y descripción de lo moral como hecho natural, propio de las ciencias, y otro que indaga acerca de la justificación y fundamentación ética, propio de la filosofía moral. Como puede apreciarse esta distinción es de una importancia crucial, habiendo sido atendida en diferentes puntos de la investigación.

Una teoría de los conceptos morales debe ser capaz de explicar cómo y por qué las ideas morales pueden influir en los seres humanos, es decir, debe ser capaz de explicar los efectos prácticos y psicológicos de dichas ideas. El estudio sobre la relación entre tales ideas y los efectos de las mismas compete a las ciencias empíricas, en la medida en que es una cuestión que puede ser explicada y descrita empíricamente. Sin embargo, no corresponde a la ciencia ofrecer respuestas a la pregunta normativa. El criterio de adecuación normativa ha de ser atendido por la filosofía moral, cuyo ejercicio trata de resolver si estamos justificados para dar peso normativo a ciertas ideas, valores, normas y principios morales. En lo que concierne al ámbito de las neurociencias, la explicación neurocientífica sobre el origen, el desarrollo y el funcionamiento de la moralidad no tiene implicaciones respecto a la validez y corrección de normas y principios morales, que han de ser fundamentados por la ética (Cortina, 2007, 2012a; Habermas, 2006; Apel, 1985; Aranguren, 1994).

Así pues, para responder a la pregunta normativa no basta con aportar una explicación neurobiológica de las conductas morales. Es preciso un criterio que pueda justificar y fundamentar las ideas morales que suscribimos. Buscamos, en palabras de la propia autora, una fuente de la normatividad, algo que respalde y fundamente la pretensión normativa de nuestras ideas acerca del deber moral. Las exigencias que imponen nuestras ideas y valores morales han de justificarse. Aquello que permita la justificación no será un objeto natural. De serlo, bastaría con el método naturalista de las

ciencias empíricas para abordar su estudio y extraer de él los códigos normativos que han de regir nuestra conducta moralmente, algunos autores defienden esta posibilidad (Mora, 2007; Gazzaniga, 2006; Churchland, 2013; Greene, 2013). Uno de los principales objetivos de nuestra investigación es someter a crítica pretensiones de este tipo, conducentes a un reduccionismo naturalista. La postura de Korsgaard es afín a la nuestra, dirige una fuerte crítica a aquellas posturas que consideran que sí es posible encontrar en la naturaleza una serie de elementos que sean intrínsecamente normativos, denomina *realismo moral sustantivo* al conjunto de tales propuestas. El bien, la justicia, la virtud, la obligación, etc. no son objetos naturales, ni puede accederse a ellos mediante la experiencia. Por ello la filosofía moral es necesaria, pudiendo abordar problemas en los que la ciencia es infecunda.

Mientras la explicación del comportamiento moral plantea una pregunta teórica en tercera persona acerca de la naturaleza de nuestra conducta, la justificación exige respuestas para una pregunta en primera persona acerca de por qué una idea o valor tiene fuerza normativa (Korsgaard, 2000; Habermas 2006; Evers, 2010; Cortina, 2012a). Así pues, la perspectiva parece ser una de las diferencias existentes entre ambos criterios. De un lado el criterio explicativo puede ser abordado de forma naturalista, atendiendo a los hechos, desde la perspectiva del observador, mientras que de otro lado el criterio normativo se resiste a una explicación de este tipo, precisando algo más que la observación, descripción y explicación de hechos, siendo necesaria la perspectiva del partícipe (Korsgaard, 2000: 27; Habermas, 2006). La confusión entre un criterio y otro tiene lugar a menudo, ya que ambos conciernen, aunque de diferente modo, a la motivación del agente para actuar correctamente.

Sin embargo, no es el mismo problema: hay una diferencia de perspectiva. Cabe la posibilidad de que una teoría que pudiera explicar por qué alguien hace lo correcto —de un modo adecuado desde la perspectiva de una tercera persona— dejara sin justificar la acción desde la perspectiva de primera persona del propio agente, y por consiguiente no apoyara sus exigencias normativas. (Korsgaard, 2000: 27)

Esa doble perspectiva ha de ser tenida en cuenta a la hora de diseñar una propuesta metodológica adecuada para la neuroética, dada su interdisciplinariedad. Ambas perspectivas han de ser integradas de forma complementaria, sin ignorar que cada una de ellas corresponde a un paradigma de investigación diferente. Los datos que la neurociencia ofrece permiten explicar y describir el comportamiento moral, pero no

justificar su validez normativa. Es ahí donde la filosofía moral ha de intervenir, tratando de dar respuesta a la pregunta normativa. Además, puede servirse de los datos aportados por la ciencia para reflexionar filosóficamente acerca de problemas morales (Moore, 1983; Teehan y diCarlo 2004). La neuroética, así, se configura como el ejercicio de la filosofía moral o ética sumado al de las neurociencias, integrados ambos en una misma especialidad.

Crítica al realismo moral

Korsgaard señala que la Filosofía Moderna ha planteado cuatro posibles respuestas para resolver la pregunta normativa: el voluntarismo, según el cual la obligación deriva del mandato de una legítima autoridad (Dios, el soberano político, etc.); el realismo, según el cual existen entidades y hechos en la naturaleza que son intrínsecamente normativos; el asentimiento reflexivo, mediante el cual podemos plantearnos si existen o no razones para aceptar una norma, idea o valor como dotados de fuerza normativa; y la apelación a la autonomía, la cual plantea que nuestra capacidad de reflexión autoconsciente nos confiere una autoridad especial sobre nosotros mismos, que permite dotar de fuerza normativa a ideas, normas y valores morales. Korsgaard reflexiona sobre cuáles de estas posturas ofrecen mejores soluciones para responder a la pregunta normativa.

Como ya hemos expuesto, lo moral acontece como un hecho natural, y en tanto que hecho natural su estudio puede ser abordado desde el método naturalista de las ciencias empíricas. Ahora bien, las posibilidades desde las cuales atender a lo moral no se agotan en la perspectiva científica. Las ciencias naturales se muestran incapaces de resolver el problema de la fundamentación y de abordar problemas concernientes a conceptos morales como el de deber, la obligación, el bien, la justicia, la virtud, etc., en la medida en que estos no son objetos naturales. Corresponderá a la filosofía moral ocuparse de estas cuestiones, pudiendo hacerse cargo del ámbito justificativo de la moral, al que corresponde ofrecer respuestas a la pregunta normativa.

El ejemplo de la teoría evolucionista muestra que una teoría podría ser adecuada para los propósitos de la explicación, y sin embargo, no dar respuesta a la pregunta normativa; y hay una razón importante para ello. La pregunta de cómo explicar el comportamiento moral es una pregunta teórica en tercera persona, una pregunta que inquiere por qué cierta especie de animales inteligentes se comporta de cierta manera, mientras que la pregunta normativa es una pregunta en

primera persona que surge para el agente moral que efectivamente debe hacer lo que la moralidad dicta. (Korsgaard, 2000: 29)

La pregunta teórica en tercera persona busca una respuesta explicativa y descriptiva, que no aportará respuestas respecto a la validez normativa de los criterios que rigen una conducta, en la medida en que sea moral. La pregunta normativa, que viene formulada en primera persona, precisa una respuesta que ofrezca una justificación de tales criterios, que vendrá marcada por un carácter evaluativo, bien desde un agente moral individual o bien desde una comunidad de agentes morales, para los cuales la obligación moral se presenta como una exigencia, al margen de que la perspectiva científica de tercera persona no pueda dar cuenta de ello. Si la perspectiva científica es insuficiente para ofrecer respuestas a las cuestiones relacionadas con la fundamentación ética es precisamente porque no existe en el mundo lo normativo como objeto natural. Tampoco existe de forma previa a la existencia de agentes morales, en algún sentido el agente moral constituye o da forma a lo normativo. Si lo normativo se encontrase en el mundo con independencia de los agentes morales, entonces sería posible que la ciencia rastrease, identificase y estudiase la fuente de la normatividad, alcanzando así conocimiento, empíricamente fundamentado, en la esfera justificativa de lo moral. Pero no sucede así. Antes de la existencia de agentes morales no puede existir por sí mismo lo bueno, lo malo, lo justo, lo injusto, lo virtuoso, lo vicioso, etc. (Moore, 1983). Solo la agencia moral abre el espacio para que estos conceptos cristalicen (Darwin, 1979; Ayala y Cela Conde, 2001; Ayala, 2006; Damasio, 2008; Waal, 1997, 2007; Brosnan y Waal, 2003; Tomasello y Vaish, 2012; Prinz, 2001; Nichols, 2005; Joyce; Hauser, 2008; Aranguren, 1994; Cortina, 2012a), los cuales no refieren a objetos del mundo, sino que corresponden a evaluaciones que solo son posibles gracias a la existencia de agentes morales. Así pues, la ciencia podrá explicar cómo un sujeto se constituye como agente moral a nivel neurobiológico y evolutivo, pero no podrá justificar los criterios normativos que regulen su conducta moralmente.

Esto no significa que lo moral caiga en el relativismo o la arbitrariedad, habrá normas y valores que podrán ser reconocidos como mejores por el agente moral o por una comunidad de agentes morales. Las diferentes teorías éticas presentan distintas propuestas en torno a la justificación y fundamentación ética.

Algunas propuestas consideran que sí existen elementos naturales que son intrínsecamente normativos, en virtud de los cuales podría determinarse qué es lo moralmente correcto, podemos encontrar estas posturas a lo largo de la historia de la filosofía, por ejemplo en algunas formas de hedonismo y utilitarismo, así como en las propuestas de algunos científicos (Spencer, 1964, 1883; Richards, 1986; Arnhart 1998) y neurocientíficos (Mora, 2007; Gazzaniga, 2006; Churchland, 2013; Greene, 2013; Casebeer, 2003). Estas posturas se sitúan dentro del llamado *realismo moral*, según refiere Korsgaard, concretamente en su vertiente naturalista, a la que llamará *realismo moral sustantivo*. Como tendremos ocasión de ver en el próximo apartado existe una vertiente procedimental del realismo que sí resulta adecuada. En cuanto al enfoque naturalista, si lo aceptásemos como válido no podríamos defender una dualidad de perspectivas metodológicas, pues la filosofía moral dejaría de ser necesaria para resolver el problema de la fundamentación, su lugar sería ocupado por una ciencia que se ocuparía de investigar los elementos normativos que podríamos encontrar en la naturaleza y que servirían para guiar la praxis humana.

Uno de los objetivos de nuestra investigación es someter a crítica estas posturas realistas, ya que incurren en un reduccionismo naturalista. Una vez descartada la validez del realismo moral será preciso reconocer que las competencias de la ciencia corresponden al ámbito descriptivo y explicativo de lo moral mientras que las de la filosofía moral corresponden al justificativo. Tal y como hemos visto en capítulos anteriores, algunos investigadores vinculados con la neuroética y ramas afines del conocimiento sostienen posturas realistas, consideran que atendiendo al cerebro humano y a la psicología podemos encontrar elementos moralmente normativos (Mora, 2007; Gazzaniga, 2006; Churchland, 2013; Greene, 2013; Casebeer, 2003). Si sus propuestas estuviesen en lo cierto la reflexión ética debería ser abandonada por estar obsoleta, y correspondería a las neurociencias describir y explicar cómo debemos comportarnos, dado que habría elementos moral e intrínsecamente normativos en la naturaleza. En cualquier caso, estos autores yerran en sus propuestas, tal y como hemos expuesto en detalle en el capítulo anterior.

Realismo sustantivo y realismo procedimental.

En el primer capítulo de las fuentes de la normatividad Korsgaard somete a crítica el voluntarismo y el realismo moral. En este apartado atenderemos a la segunda de las

críticas, ya que resulta especialmente relevante para nuestra investigación. En cuanto al voluntarismo, volveremos sobre él más adelante. El realismo moral sustantivo es naturalista, sostiene que hay exigencias morales que tendrán fuerza normativa si son verdaderas, y serán verdaderas en la medida en que describan correctamente entidades o hechos intrínsecamente normativos. Además, sostiene, por este motivo, que tal tipo de entidades o hechos existen. Nuestra pretensión en este apartado será poner en tela de juicio estas pretensiones. Los valores, las obligaciones o las razones, moralmente entendidos, son algunos de esos elementos que el realista considera como existentes en el mundo y normativos en sí mismos. Korsgaard identifica estas posturas realistas con lo que llama *realismo sustantivo*, en el cual se centrará nuestra crítica, y lo diferencia de una clase de realismo que podría calificarse de *procedimental*, que sería aceptable en los términos expuestos por la autora. Atendamos a esta distinción.

El realismo procedimental no es naturalista, sostiene que hay respuestas a las preguntas morales, pero no en virtud de que existan entidades o hechos naturales intrínsecamente normativos, sino porque hay modos correctos e incorrectos de responderlas. Responder adecuadamente a las preguntas morales es posible mediante ciertos procedimientos, a pesar de que no haya elementos intrínsecamente normativos en la naturaleza. Propuestas como las de la ética dialógica (Apel; Habermas; Cortina) o el equilibrio reflexivo de Rawls (Rawls, 2013) plantean una metodología procedimental de este tipo. Lo que el realismo procedimental sugiere, en definitiva, es que existen procedimientos que resultan adecuados para responder preguntas morales, en virtud de ello será posible establecer normas morales y hablar de conceptos morales como lo bueno, lo correcto y lo justo, sin hacerlo arbitrariamente. El realista procedimental sería realista en tanto que habría enunciados morales verdaderos, “los enunciados que implican conceptos morales serán verdaderos cuando esos conceptos se apliquen correctamente” (Korsgaard, 2000: 52). En este punto es conveniente hacer una observación, parece más adecuado hablar de validez, como proponen otros autores (Habermas; Cortina), que de verdad moral, en los próximos capítulos atenderemos a estos autores para profundizar en esta cuestión y ofrecer una terminología adecuada. El realismo procedimental de Korsgaard, en los términos descritos, se plantea de forma afín al cognitivismo en la medida en que ambos plantean la posibilidad de argumentar sobre el deber moral.

El realista sustantivo también considera que pueden ofrecerse respuestas verdaderas a las preguntas morales, la diferencia radica en que achaca esa posibilidad a la presunta existencia en la naturaleza de entidades y elementos intrínsecamente normativos. Por su parte, el realista procedimental piensa que tales respuestas son posibles gracias a procedimientos que hacen adecuada la respuesta. Respecto a ello, el realista sustantivo cree que cualquier procedimiento sería solo un modo de informarse sobre hechos naturales, intrínsecamente normativos, que existirían con independencia del propio procedimiento, y que serían los que permitirían ofrecer las respuestas al problema de la fundamentación y la justificación normativas.

El realismo sustantivo versa, por tanto, sobre qué clase de disciplina es la ética. Si la ética se ocupa de conocer una parte del mundo, la parte normativa, que existiría como hecho natural, entonces sería una ciencia de lo normativo, dejando ya de ser filosofía moral (Spencer, 1964, 1883; Richards, 1986; Arnhart 1998; Singer, 1981; Casebeer, 2003: 34; Flanagan; Cartwright, 2000: 328; McShea y McShea, 1990, 1999; Churchland, 2013; Curry, 2006: 235, 237-238; Mora, 2007; Gazzaniga, 2006; Greene, 2013).

El realismo sustantivo entiende los procedimientos para responder preguntas normativas como maneras de *informarse* acerca de determinada parte del mundo: la parte normativa. En esta medida, el realismo moral sustantivo no se distingue por su posición sobre qué clase de verdades hay, sino por su posición acerca de qué clase de disciplina es la ética. Entiende la ética como una rama del conocimiento, conocimiento de la parte normativa del mundo. (Korsgaard, 2000: 53)

Según el realismo sustantivo, pues, la ética es en realidad una disciplina teórica o epistemológica. Cuando planteamos preguntas éticas o, de manera más general, preguntas normativas prácticas, estamos tratando de averiguar algo acerca del mundo. El mundo contiene un ámbito de entidades o verdades inherentemente normativas cuya existencia hemos advertido, y a la ética, o a la filosofía práctica en general, le corresponde investigar más a fondo acerca de ellas, aprender sobre ellas de una manera más sistemática. (Korsgaard, 2000)

El realismo moral sustantivo resulta problemático, quedando refutado en la medida en que la naturaleza no es moralmente normativa por sí misma, tal y como tratamos de mostrar en la presente investigación. Por ello, defender que la ética es, o debe ser, una ciencia natural que investiga hechos y entidades intrínsecamente normativos conduce inevitablemente a un reduccionismo naturalista y cientificista, que no resiste la crítica (Buchanan, 2009; Apel, 1985; Habermas, 2006; Aranguren, 1994; Gould, 1999;

Hume 1988, 1953; Moore, 1983; Salles, 2013, 2016, 2018; Evers, 2010; Cartwright, 2010; Teehan y diCarlo, 2004; García-Marzá, 2016; Conill, 2012).

Entidades y hechos intrínsecamente normativos según el realismo moral.

Aunque no estemos de acuerdo con la postura del realismo sustantivo, cabe al menos preguntarse qué clase de entidades o hechos son esos que el realista considera como intrínsecamente normativos. Existen diferentes propuestas desde el realismo moral, atenderemos a algunas de ellas.

Podría pensarse que esas entidades intrínsecamente normativas precisan de una percepción especial, no serían hechos convencionalmente tangibles, sino que se precisaría tener ciertas capacidades perceptivas especiales. Las formas de Platón y los valores intrínsecos no naturales de Moore son propuestas que encarnan ejemplos de ello, se comprometen de uno u otro modo con la existencia de cierta clase de objetos metafísicos. Sin embargo, otras propuestas pretenden situar la fuente de la normatividad en elementos accesibles a la experiencia, y no a la intuición intelectual. Tal es la propuesta de Thomas Nagel (1985) y de algunos neuroeticistas como Gazzaniga (2006) y Mora (2007), entre otros.

Lo que Nagel plantea es que existen ciertos intereses naturales humanos, como por ejemplo el de obtener placer y evitar dolor, los cuales parecen tener cierto carácter normativo. El dolor y el placer, los deseos y las aversiones, y los intereses naturales, parecen todos ellos constituir razones para la acción, razones que tendrían un carácter normativo. La evitación del dolor y la promoción del placer parecen buenos criterios para reflexionar en torno a la adecuación moral de una norma o acción, contemplar tales intereses es un criterio relevante para que la reflexión ética y sus conclusiones no sean arbitrarias. Ahora bien, eso no significa, en primer lugar, que sean los únicos criterios posibles, ni tampoco, en segundo lugar, que el placer y el dolor sean normativos en sí mismos.

Provocar placer y dolor innecesaria y deliberadamente puede condenarse moralmente, pero no en virtud de que el placer y el dolor sean normativos en sí mismos. Corresponde a la reflexión y la argumentación ética evaluar moralmente la acción que produce placer o dolor. Todo ello cae en una esfera reflexiva, evaluativa y valorativa

propia del agente moral o de una comunidad de ellos, siendo estos quienes proyectan la normatividad sobre el mundo, y no al revés.

El placer y el dolor tienen una función biológica que permite regular la interacción de determinados seres con su medio, son útiles, necesarios e inevitables. Pero solo el agente moral puede tomarlos como un factor relevante a la hora de reflexionar sobre la justificación moral de las acciones que los producen. Puede razonarse por qué un actor moral no debe infligir daño deliberada e innecesariamente, pero el dolor y el placer no son normativos de forma intrínseca. En lugar de ello, se ha de atender a la valoración del agente moral, que considera el dolor y el placer como algo negativo y positivo, respectivamente, en determinados contextos. Que un ser sintiente sufra dolor o placer no es bueno ni malo moralmente mientras que no haya agentes morales, por un lado, estos pueden evaluar moralmente la acción y, por otro, es a estos a quienes se les puede imputar la responsabilidad respecto a aquellas acciones que producen ese placer o dolor.

La necesidad de la filosofía moral

Si nos preguntamos por qué tenemos el deber moral de hacer algo entraremos en una serie de preguntas que remitirán a otras, iniciando un regreso *ad infinitum* en la búsqueda de una fundamentación ética (Korsgaard, 2000: 49; Kant, 1992, 2000; Cortina, 2007). La razón busca lo incondicionado para poner fin a esa regresión al infinito. Ese elemento incondicionado ha de ser intrínsecamente normativo, de tal manera que “la respuesta incondicional deberá hacer imposible, innecesario o incoherente volver a preguntar por qué” (Korsgaard, 2000: 49). La pretensión de ontologizar naturalistamente lo normativo, que pertenece a la esfera práctica, es precisamente el error cometido por el realista moral sustantivo al defender que hay elementos intrínsecamente normativos en el mundo.

En cuanto a la regresión al infinito que supone la búsqueda de lo incondicionado, la estrategia del realista moral sustantivo es poner fin a esa regresión por decreto, declara la existencia de objetos naturales que serían normativos por sí mismos. De este modo, el objeto que permite establecer criterios normativos sería natural, motivo por el cual la ética debería constituirse como ciencia empírica y no como reflexión filosófica. En oposición a esta idea, Korsgaard expone que los *objetos normativos* son ideas que la razón produce

en su búsqueda de las fuentes de la normatividad, por ello resulta más apropiado referirse a ellos como *conceptos normativos*.

Cabría preguntarse, entonces, por qué tenemos conceptos normativos si no existe ningún objeto normativo en la naturaleza. ¿Qué empuja al ser humano a crear conceptos que le ayuden a guiar sus acciones en sentido moral? ¿Por qué nuestra especie tiene esa clase de conceptos normativos que no refieren a ningún objeto natural? La respuesta a esta pregunta sí compete a la ciencia, ya que las capacidades y facultades psicobiológicas que permiten tener conceptos de este tipo pueden ser explicadas y descritas desde estudios empíricos. Como vimos en capítulos anteriores (Darwin, 1979; Ayala y Cela Conde, 2001; Ayala, 2006; Damasio, 2008; 013; Waal, 1997, 2007; Brosnan y Waal, 2003; Tomasello y Vaish, 2012; Prinz, 2001; Nichols, 2005; Joyce; Hauser, 2008; Cortina, 2012a; Aranguren, 1994), y como Korsgaard expone, el ser humano es un animal social, con ciertas capacidades y facultades cognitivas que le han posibilitado y le han empujado a la reflexión sobre los motivos de su propia praxis en la esfera social, siendo capaz de preguntarse qué acción es mejor en sentido moral y por qué. Todo ello abre el espacio de la ética, en el cual se busca un criterio para la acción que no sea meramente biológico, que no atienda a lo que *es*, sino a lo que *debe ser*. Eso no significa que exista en la naturaleza el *deber moral*, significa que el ser humano es capaz de concebir conceptualmente ese deber, cuya justificación se plantea como un problema, al cual ha de ofrecer una solución. Los conceptos normativos son uno de los resultados de esta capacidad y de la búsqueda de respuestas a ese problema.

Los conceptos normativos existen porque los seres humanos tenemos problemas normativos, y tenemos problemas normativos porque somos animales racionales autoconscientes, capaces de reflexionar sobre lo que debemos creer y hacer. A eso se debe que la pregunta normativa pueda siquiera plantearse: porque aun cuando nos inclinemos a creer que algo es correcto y hasta cierto punto nos sintamos movidos a hacerlo, *siempre* podemos preguntar: ¿Pero es realmente cierto? y ¿De verdad tengo que hacerlo? (Korsgaard, 2000: 65)

Los motivos naturales en virtud de los cuales tenemos problemas y conceptos normativos pueden ser estudiados por la ciencia desde una perspectiva explicativa y descriptiva. Sin embargo, cómo resolver esos problemas y cómo aplicar y gestionar los conceptos normativos es algo respecto a lo cual la ciencia no tiene competencias, será la filosofía moral la que se ocupe de esa tarea. Por ello, la dualidad de perspectivas metodológicas propia de la neuroética, científica y filosófica, no presenta problemas de

complementariedad, ambas pueden abordar la cuestión moral, siempre que cada una de ellas se limite a ámbitos que les competen, a saber, el explicativo y el justificativo, respectivamente.

Una neuroética no reduccionista

Las diversas propuestas del realismo moral que pretenden fundamentar una ética a través de la ciencia (Spencer, 1964, 1883; Richards, 1986; Arnhart, 1998; Singer, 1981; Flanagan; Cartwright, 2000: 328; McShea y McShea 1990, 1999; Curry, 2006: 235, 237-238), mediante el método empírico de esta, atendiendo a hechos y entidades naturales que consideran intrínsecamente normativos, en las cuales se enmarcan algunas propuestas neurocientíficas (Casebeer 2003; Churchland 2013; Mora, 2007; Gazzaniga, 2006; Greene, 2013), son una empresa condenada al fracaso. La filosofía moral tiene mucho que decir y ofrecer como disciplina destinada a orientar la praxis. No es posible desprendernos de una dualidad de perspectivas metodológicas en el seno de la neuroética, la ciencia no puede hacerse cargo del ámbito justificativo de la moral, pues la perspectiva científica está constitutivamente imposibilitada para ejercer una labor de fundamentación ética. La neuroética no es ni puede ser una disciplina exclusivamente científica, sino que es un campo interdisciplinar donde diversas ramas de las ciencias empíricas interactúan con la filosofía moral. Entre esas ramas de la ciencia se encuentran la biología, la etología, la sociología y la psicología, que cada vez van adoptando más competencias en el ámbito neurocientífico. En el campo de la neuroética, la interacción entre neurociencias y filosofía moral se ha de producir de tal modo que sus respectivas perspectivas metodológicas puedan funcionar complementariamente sin necesidad y sin posibilidad de reducir una a la otra.

Existen propuestas en neuroética que reconocen la importancia y la necesidad de hacer complementarias las perspectivas metodológicas de las que hablamos (Habermas, 2006; Cortina, 2012a; Salles, 2013, 2016, 2018; Evers, 2010). Tales propuestas evitan el reduccionismo, no tienen pretensión de deducir unas normas morales a partir del estudio del cerebro, ni pretenden que la ciencia sea capaz de descubrir elementos intrínsecamente normativos. Estas propuestas son verdaderamente valiosas y representan ejemplos donde toma forma un método científico-filosófico adecuado para la neuroética.

Recordemos la distinción entre estructuras y contenidos morales (Aranguren, 1994; Hauser, 2008) y entre bases y fundamentos de la moralidad (Cortina, 2012a: 94-96). La estructura moral está formada por la base natural de las conductas morales, que viene dada neurobiológicamente y que condiciona y orienta la conducta de forma probabilística, pero no la determina. Los contenidos son contingentes e históricos y se asientan sobre esas estructuras morales, su fundamentación no dependerá de estas, sino de la ética o filosofía moral. Tales estructuras componen las bases psicobiológicas de la moralidad, cuyo estudio compete a la ciencia, mientras que los fundamentos éticos de los contenidos morales tendrán como objetivo justificar estos, esta tarea compete a la ética. Esta doble distinción es un indicador de la necesidad de una metodología dual en neuroética, en la que se integre la ciencia y la filosofía moral.

Otros autores, sin embargo, pretenden también diseñar una ética universal desde las neurociencias, la psicología evolutiva y la sociobiología, es decir, recurrir a sus métodos empíricos, pero se distinguen de los primeros por tres razones al menos: no pretenden llegar con su investigación a descubrir los *contenidos* de una ética universal, sino sólo a descubrir una estructura moral universal que se modula de forma diferente en las distintas culturas; sus propuestas de neuroética están mucho más elaboradas que las del grupo anterior y permiten diseñar un cierto marco teórico; y, por último, no rehúsan la ayuda de la filosofía, y no sólo porque dialogan con algunas teorías éticas relevantes, sino también porque recurren a algún método filosófico para llevar a cabo su tarea. (Cortina, 2012a: 57).

Así pues, en la tarea de fundamentación ética desempeñada por la filosofía moral, esta podrá valerse de los datos que la neurociencia aporte, siempre que sean relevantes para la reflexión y la argumentación (Moore, 1983; Teehan y diCarlo 2004). Sin embargo, aunque la ciencia pueda informar a la ética en algunos casos, los datos que aporte no podrán ser tomados como criterios suficientes para la justificación, ya que esta labor es una competencia propia de la filosofía moral.

Reflexividad y normatividad

Hasta aquí hemos atendido dos cuestiones que Korsgaard trata y que resultan especialmente relevantes para abordar la cuestión del método en neuroética. La primera de ellas es la diferenciación entre el ámbito explicativo y el justificativo en la moral. La segunda de ellas es la crítica al realismo moral sustantivo, por cometer reduccionismo naturalista. Estas cuestiones, junto a una crítica al voluntarismo, en tanto que propuesta

sobre el origen de lo normativo, constituyen para Korsgaard el contexto filosófico a partir del cual desarrollará su propuesta.

Tras exponer en *Las fuentes de la normatividad* sus consideraciones sobre la cuestión de la fundamentación ética, en la conferencia titulada “La pregunta normativa”, Korsgaard pasará a bosquejar su postura acerca de cómo responder a esa pregunta, tarea que llevará a cabo en tres conferencias que llevan por título “El asentimiento reflexivo”, “La autoridad de la reflexión” y “El origen del valor y el alcance de la obligación”. En ellas expondrá desde una postura heredera de Kant cómo la reflexión en torno a nuestros impulsos para la acción permite asentir ante estos si los consideramos adecuados, siendo así la autonomía la fuente de la normatividad.

Además de la crítica al voluntarismo y al realismo, Korsgaard somete a revisión una postura que denomina *asentimiento reflexivo*, en el cual puede situarse el pensamiento kantiano. El asentimiento reflexivo presenta dos momentos. El primero de ellos consiste en indagar acerca de cuáles son los fundamentos de la naturaleza humana, con el objetivo de conocer qué nos mueve a la acción moral, por qué empleamos conceptos morales, por qué nos sentimos obligados por ellos, etc. Este momento se ocupa del ámbito explicativo de la moral. Una vez se disponga de una explicación de nuestra naturaleza moral deberemos someter a crítica las disposiciones conductuales hacia las que nos inclina esa naturaleza, ese es el segundo momento, en él se plantea la pregunta normativa y se busca una justificación de lo moral.

Independientemente de las predisposiciones e inclinaciones que nuestra biología o nuestra cultura nos impongan, siempre podemos someter a juicio las creencias e impulsos que influyen en nuestras acciones (Darwin, 1979; Ayala y Cela Conde, 2006; Ayala, 2001; Damasio, 2008; Waal, 1997, 2007; Brosnan y Waal, 2003; Tomasello y Vaish, 2012; Prinz, 2001; Nichols, 2005; Joyce; Hauser, 2008; Aranguren, 1994; Cortina, 2012a; Evers, 2010). Solo la aprobación o el asentimiento ante esas acciones, creencias y motivos, llevados a cabo tras un ejercicio reflexivo adecuado, otorgan estatus normativo a los mismos (Kant, 1992, 2000; Cortina, 2007; Apel, 1985; Habermas, 2010, 2006; Haidt, 2001, 2013; Darwin, 1979: 94, Korsgaard, 2000). Es decir, solo justificándolos reflexivamente podemos asentir ante ellos como códigos de conducta moralmente justificados. La búsqueda de fundamentación ética participa de esta idea, tratando de buscar las razones que permitan asentir ante ciertas ideas y valores, rechazando otros. Así

pues, Korsgaard plantea el asentimiento reflexivo como un método filosófico que permite establecer unas normas morales, justificándolas racionalmente.

Si tras reflexionar sobre la verdadera teoría moral descubrimos que seguimos inclinados a asentir a las exigencias de la moralidad, entonces la moralidad será normativa. A esta manera de establecer la normatividad la denomino el método del «asentimiento reflexivo». (Korsgaard, 2000: 70)

No ha de pasar desapercibido el hecho de que Korsgaard habla de *establecer* la normatividad y no de descubrirla como algo previamente existente. Esta cuestión será clave para la respuesta que más adelante la autora plantea a la pregunta normativa, así como para el tratamiento que haremos de la cuestión en los siguientes capítulos. Korsgaard, mostrando su afinidad con Kant, considera que la fuente de la normatividad es la autonomía del individuo, mediante el ejercicio del asentimiento reflexivo, en una forma de voluntarismo autónomo que permite la producción de máximas que serán entidades normativas, constituyendo así una forma de realismo procedimental.

En su análisis atiende a las propuestas de tres autores muy diferentes, argumentando que todos ellos suscriben el asentimiento reflexivo de uno u otro modo. Estos son David Hume, con una propuesta relativa a la cuestión del sentimiento, Bernard Williams, con una propuesta de corte aristotélico, en defensa de que la normatividad ha de ser compatible con el desarrollo humano, y John Stuart Mill, quien plantea cierto realismo en relación al principio de utilidad, pero considerándolo insuficiente, por lo cual apela a la reflexión y al asentimiento para establecer la normatividad. No es necesario profundizar y exponer cada una de estas propuestas, sino más bien destacar la presencia del asentimiento reflexivo en todas ellas y cómo desembocan en una posición de corte kantiano, según defiende Korsgaard. En lo que difieren estos tres autores es en qué elementos son relevantes para la reflexión, lo cual puede conducir a diferentes resultados. Pero todos coinciden, asegura la autora, en que la reflexión y el asentimiento son necesarios para establecer la normatividad, que no se halla en la naturaleza, sino que viene proyectada por el agente moral autorreflexivo. Hume considera que no hay nada en una acción que la haga abominable o viciosa, sino que es nuestra evaluación como agentes morales la que proyecta sobre la acción la abominabilidad y el vicio (Hume, 1988). Del mismo modo Bernard Williams (Williams, 1972) considera que es el agente moral quien proyecta el valor ético en el mundo a través de sus creencias, convicciones y evaluaciones

morales. La autorreflexión deberá determinar cuándo tales proyecciones son susceptibles de asentimiento.

Si el asentimiento reflexivo de nuestras disposiciones es lo que establece su normatividad, lo que necesitamos a fin de establecer la normatividad de nuestros motivos e inclinaciones más particulares es el asentimiento reflexivo de estos. Después de todo, en eso radica todo el sentido de emplear el método del asentimiento reflexivo para justificar la moralidad: suponemos que cuando reflexionamos sobre las cosas a las que descubrimos estar inclinados, podemos entonces aceptar o rechazar la autoridad que esas inclinaciones pretenden tener sobre nuestra conducta y actuar en consecuencia. (Korsgaard, 2000: 116)

Hume y Williams utilizan el asentimiento reflexivo como método para establecer la normatividad de nuestras disposiciones y sentimientos morales, dadas por nuestra naturaleza. Pero Korsgaard señala que ese método “es precisamente el proceso mental que, según Kant, caracteriza las deliberaciones del agente moral autónomo” (Korsgaard, 2000: 116). Someter a reflexión cada impulso y motivo de la acción permite determinar si es una razón para la acción. Así el asentimiento reflexivo de Hume y Williams desemboca en realidad en la propuesta kantiana de la autonomía como fuente de la normatividad. Bajo el supuesto de que una razón es intrínsecamente normativa, uno de los modos de analizar si un motivo o impulso es una razón será preguntarnos si podemos querer que la máxima de actuar conforme a ese motivo se convierta en ley universal. Así, el asentimiento reflexivo no es solo una forma de asentir o disentir respecto a disposiciones y sentimientos morales, sino que es el modo en el que se someten a prueba los motivos de la acción para ver si son razones. Aquí encontramos el punto en el que descansa el error de algunos neuroeticistas como Greene y Haidt, al considerar que la intuición es la base de la conducta moral, mientras que la razón solo es una justificación *post hoc* de convicciones morales originadas en la intuición (Greene; Haidt, 2001, 2013). Es preciso matizar que el intuicionismo social de Haidt, a diferencia de la propuesta de Greene, considera que la razón desempeña un papel de justificación en algunos casos, pero según plantea el autor este papel crítico de la razón solo intervendría en un número muy restringido de casos, protagonizados por individuos especialmente dotados para ello, como los filósofos. Esto permite someter a crítica la idea, esgrimida por Greene, de que el deontologismo es una racionalización de motivos intuitivos más que un medio legítimo de fundamentación de normas morales, tal y como expusimos en el Capítulo 3.

[...] la prueba del asentimiento reflexivo es empleada por agentes morales para establecer la normatividad de todos sus motivos e inclinaciones particulares. La prueba del asentimiento reflexivo no es, pues, solamente una manera de justificar la moralidad, sino que es la mismísima moralidad. (Korsgaard, 2000: 116)

Todo impulso debe pasar una prueba crítica para poder ser aceptado éticamente como razón para la acción, esa prueba es el asentimiento reflexivo, que constituirá la moralidad misma en la medida en que esa prueba no descubre, sino que establece lo normativo. Con esta cuestión se adentra Korsgaard en la tercera de sus conferencias, en la cual plantea principalmente dos conclusiones. La primera de ellas, a la que ya hemos apuntado, es que la autonomía es la fuente de la obligación, especialmente en relación a uno mismo; la segunda es que tenemos obligaciones morales hacia la humanidad como tal.

La normatividad se presenta al ser humano como algo problemático porque la mente humana es autoconsciente, la capacidad de dirigir nuestra atención a nuestras propias actividades mentales, pudiendo cuestionarnos por qué un impulso es moralmente más deseable que otro, abre el espacio de la reflexión ética (Darwin, 1979; Ayala y Cella Conde, 2001; Ayala, 2006; Damasio, 2008; 013; Waal, 1997, 2007; Brosnan y Waal, 2003; Tomasello y Vaish, 2012; Prinz, 2001; Nichols, 2005; Joyce; Hauser, 2008; Aranguren, 1994; Cortina, 2012a; Evers, 2010). La capacidad de distanciarnos de esos impulsos y preguntarnos por qué hemos de actuar conforme a ellos, si existen alternativas y cuál es mejor, sumados a nuestra naturaleza social, abre todo un campo de reflexión en torno a la praxis que versa no solamente sobre qué nos mueve a la acción, sino por qué y cómo debería movernos. La mente reflexiva necesita una razón, necesita poder asentir ante los impulsos y disposiciones naturales para considerar que tienen fuerza moralmente normativa. La constitución natural del agente moral compone las condiciones de posibilidad de la moralidad, cuya aparición genera el problema de cómo dirigir la conducta para actuar de forma ética. El agente moral es en este sentido la fuente de la normatividad. Korsgaard dirá que lo es en tanto que legislador autónomo, mediante asentimiento reflexivo.

La autora sostiene que las razones son resultado de un éxito reflexivo, una razón para actuar es un impulso que cuenta con la bendición del agente moral tras someterse a reflexión. Otra posibilidad que cabe añadir es que la razón puede generar un nuevo

impulso para la acción cuando los impulsos preexistentes no resisten la crítica racional. Del mismo modo que las razones surgen de un asentimiento reflexivo, Korsgaard sostiene que la obligación surge del rechazo reflexivo. La negativa del agente moral para actuar de cierto modo, resultado de reflexionar en torno a un impulso o una acción, produce un rechazo reflexivo que, al tiempo que obliga a no actuar de cierto modo, obliga a actuar de otro.

Cada agente moral, además, tiene una identidad práctica, conformada por la comprensión de sí mismo, cómo uno se valora, cómo uno encuentra sentido a la vida y por qué considera que merece la pena emprender unas acciones y no otras. Nuestro asentimiento y nuestro rechazo reflexivo vendrán íntimamente relacionados con esa identidad práctica, que estará siempre compuesta racional y emocionalmente, y sujeta a modificaciones. Vemos aquí una dimensión emotivo-racional en el agente moral, que sintoniza con la propuesta de una ética cordial (Cortina, 2009) y con algunas propuestas de neuroeticistas afines (Evers, 2010). La obligación deriva de la negativa a actuar de modo tal que se diluya nuestra identidad práctica, mientras que todo asentimiento promueve esta identidad convirtiendo en razones aquellos impulsos compatibles con esta.

Korsgaard comienza su obra con una crítica al voluntarismo y otra al realismo, después lleva a cabo una revisión de lo que denomina asentimiento reflexivo, y finalmente propone una apelación a la autonomía como fuente de la normatividad. La relación entre el asentimiento reflexivo y la autonomía como fuente de la normatividad ya ha sido puesta de relieve. Korsgaard establece también un vínculo con el voluntarismo y con el realismo, los cuales ha sometido a crítica al principio de su obra, reconociendo en ellos algunas formas que resultan adecuadas para atender a la cuestión de la fundamentación.

Respecto al voluntarismo, este atribuye a un legislador potestad normativa, siendo este legislador la fuente de la normatividad. Dice nuestra autora que “el yo actuante concede al yo pensante el derecho de gobernar” y este “procura gobernar lo mejor que puede” (Korsgaard, 2000: 133). Así la conciencia humana se estructura en una relación de autoridad del sujeto consigo mismo, siendo esa autoridad una fuente de obligación. Lo que Korsgaard está diciendo es que la fuente de la normatividad recae en última instancia en un legislador que tiene autoridad, el yo reflexivo.

En cuanto al realismo, lo que Korsgaard hará será defender un realismo procedimental. Toda la crítica que hicimos respecto al realismo sustantivo, así como la relación de esa crítica con la neuroética, sigue vigente. Para defender un realismo procedimental, la autora busca el procedimiento adecuado para la producción de máximas morales. Señala que las buenas máximas son un producto de la voluntad legislativa de los individuos humanos, el descubrimiento de una máxima buena no la torna en ley hasta que la voluntad legislativa no entra en acción. En ese sentido los humanos somos creadores de valores. De ahí que el realismo que Korsgaard defiende sea procedimental, “los valores se construyen mediante un procedimiento, el procedimiento de hacer leyes para nosotros mismos” (Korsgaard, 2000: 143). Uno de los ejemplos de procedimentalismo que cita es el de Rawls en *Liberalismo Político* (2013). Nosotros atenderemos, en los próximos capítulos, a las éticas dialógicas, que también son procedimentales en este sentido.

Así Korsgaard reúne en su propuesta las posturas del voluntarismo, del realismo —como realismo procedimental—, del asentimiento reflexivo y de la apelación a la autonomía kantiana. La crítica al realismo sustantivo, en tanto que reduccionista, se mantiene.

Dado que la existencia de la agencia moral es lo único que permite una apertura para el espacio de lo ético y lo normativo, la propuesta de Korsgaard resulta muy atractiva, explica cómo desde la agencia moral puede establecerse lo normativo, y además aporta una propuesta para este establecimiento de normas. Así, lo normativo queda sujeto a la contingencia, pero no a la arbitrariedad, puesto que sería resultado de una reflexión racional, bien sea privada o pública, que trataría de dar una justificación. De esta manera, lo normativo no es algo realizado, sino que se constituye en continua argumentación y reflexión, siempre revisable, ya sea monológicamente o dialógicamente.

En la cuarta conferencia, titulada “El origen del valor y el alcance de la obligación”, Korsgaard aborda tres cuestiones. La primera plantea la cuestión de si valorar nuestra propia humanidad compromete o no valorar la de otros. La segunda concierne a la cuestión de si la valoración de la humanidad en las teorías kantianas excluye o no la posibilidad de otorgar un estatus o categoría moral a otros animales. Por último, atenderá a la cuestión del escepticismo normativo.

La respuesta de Korsgaard a la primera de estas cuestiones es que solo una concepción errónea de las razones, entendidas como entidades privadas, puede conducir a la conclusión de que valorar nuestra propia humanidad no nos compromete a valorar la de los demás. Si considero mi humanidad como fuente de valor, y teniendo en cuenta que mi humanidad es análoga a la de otros seres humanos, deberé reconocer el valor de los otros seres humanos en tanto que humanos, siempre que reconozca el mío en tanto tal. Debo valorar a los otros y a las cosas que ellos valoran exactamente en el sentido en que me valoro a mí mismo y las cosas que yo valoro.

Korsgaard trata de mostrar que las razones no son privadas, sino públicas en esencia. Si el realismo sustantivo estuviese en lo cierto, la publicidad de las razones vendría dada por la objetividad de ciertas características y hechos naturales, pero el realismo sustantivo se equivoca, así que hay que buscar otro modo de mostrar la publicidad de las razones. La opción es que Korsgaard sugiere es que las razones son inherentemente compartibles, aún si parten o se dan originariamente de forma privada, pueden ser compartidas con los demás.

El carácter público de las razones se crea, efectivamente, a partir del intercambio recíproco de las razones de los individuos, a partir del acto de compartirlas. Sin embargo, esto reconoce la cuestión señalada por las críticas anteriores: si estas razones fueran esencialmente privadas, sería imposible intercambiarlas o compartirlas. Su condición de privadas (su privacidad) debe de ser, entonces, incidental o efímera; las razones tienen que ser inherentemente compartibles. (Korsgaard, 2007: 170)

Al hacer que otro tome en consideración los términos de mi reflexión, mis impulsos y las razones por las que asiento ante ellos, fuerzo al otro a reflexionar. Este, a su vez, me fuerza a argumentar correctamente. La puesta en público de las razones pasa así a convertirse en una suerte de reflexión pública en la que el asentimiento se produce en el intercambio público de argumentos, en el diálogo. La cercanía de esta postura con la de las éticas dialógicas es evidente, pero también es afín a otros procedimentalismos normativos como el de Rawls (1995, 2013). Dado que Korsgaard defiende una postura de corte kantiano, no es de extrañar que al tomar las razones como algo público su postura desemboque en un planteamiento afín al procedimentalismo dialógico y al rawlsiano, que son también herederos de Kant y que tienen en cuenta el ámbito público de la razón. Además, en la medida en que Korsgaard habla del carácter público no solo de las razones, sino también de los impulsos, los rasgos emocionales de la argumentación se sitúan como

un elemento relevante, que encuentra en la ética de la razón cordial una de las propuestas que mejor visibilizan y sistematizan su relevancia y su papel en el desarrollo de una propuesta dialógica (Cortina, 2007, 2009).

En lo que concierne a la cuestión de los animales no humanos, Korsgaard presta especial atención al hecho de que son seres que sufren dolor. El dolor es un mecanismo que impulsa al individuo a cambiar de situación, es una alarma que regula la evitación de peligros potenciales. Asimismo, el dolor se experimenta de forma consciente, asegura, aunque no sea de forma autoconsciente. Todo ello, cabe señalar, no significa que sea intrínsecamente normativo, en ausencia de agentes morales el dolor es un hecho de la naturaleza sin carga moral alguna. Es el agente moral quien reconoce el dolor como algo que permite asentir reflexivamente a aquellos impulsos que permiten reducirlo, evitarlo y paliarlo, tanto en nuestro propio cuerpo como en el de otros. Según la teoría kantiana tal y como la ha expuesto Korsgaard, las expresiones de dolor en el caso de los humanos suelen y pueden tomar la forma de una puesta en público de razones, que obliga a los otros a atender a estas razones tanto como atienden a las suyas. Podemos considerar que lo mismo sucedería en el caso de los animales cuando estos sufren y expresan dolor, pero para ello hace falta considerar que los animales tienen razones, aunque no sea de forma autoconsciente y reflexiva. Esa será la propuesta de Korsgaard. Considera que el dolor es la percepción de una razón, el dolor es el rechazo no reflexivo de una amenaza a la identidad propia —que también el animal no humano tiene en tanto que busca su propia conservación—. El dolor es, pues, la percepción de una razón, y en tanto que podamos atribuirlo a un animal, por ejemplo interpretando sus expresiones de dolor, nos veremos forzados a tomar en consideración sus razones, aunque no sean autoconscientes, de igual modo que nos vemos forzados a tomar en consideración las razones de otros seres humanos al ser expuestas públicamente. La condición no autorreflexiva de los animales en relación a sus razones y a las de los demás hace que estos sean beneficiarios de nuestra consideración moral, pero no deudores, tal como pasa en algunos casos de humanos, cuyo uso de ciertas facultades no está desarrollado o presenta alguna patología. El estatus o valor moral de un ser puede imponer exigencias morales a los demás sin que esto signifique que dicho ser tenga obligaciones morales. Para tener deberes morales se precisa ser un agente moral, lo cual implica tener las facultades necesarias para comprender y asumir las razones de los otros y el sentido en que estamos obligados moralmente hacia

ellos. Sin embargo, para ser susceptible de consideración moral no resulta necesario ser un agente moral, por los motivos expuestos.

Esta cuarta conferencia cierra con una afirmación que considero problemática y que quiero matizar. Korsgaard acaba afirmando que el realismo moral está en lo cierto.

[...] es cierto que hay tipos extraños de entidades, y que conocerlas no se parece a ninguna otra cosa. Pero eso no significa que no existan. [...] pues el hecho más conocido de la vida humana es que el mundo contiene entidades que nos dicen qué hacer y que nos hacen hacerlo. Se trata de las personas, y de los otros animales. (Korsgaard, 2000: 206)

Según expone Korsgaard en este fragmento, las personas y los animales son intrínsecamente normativos, ya que son entidades que ofrecen razones capaces de forzar al agente moral a reflexionar en torno a ellas, asintiendo ante ciertas normas cuando el análisis racional lo considere oportuno. Es preciso señalar, en relación a ello, que la fuente de la normatividad no está en quién ofrece razones, ya que esto puede hacerse de forma no reflexiva, como sucede en el caso de los animales. Solo el agente moral puede evaluar moralmente los hechos, las situaciones, los impulsos, etc., de tal forma que pueda asentir reflexivamente y establecer así lo normativo. Las personas y los otros animales pueden influirnos para actuar de cierta manera, pero no porque sean normativos en sí mismos, sino porque dirigen nuestra propia reflexión en cierta dirección, nos aportan razones, que en ocasiones tienen de forma no reflexiva, siendo el agente moral el único que puede asentir o disentir ante tales razones, estableciendo así lo normativo. El agente moral, en ese sentido, es la fuente de la normatividad. La normatividad no es otra cosa que las conclusiones que los agentes morales extraen en sus reflexiones y acuerdos sobre lo normativo.

El asentimiento reflexivo en el método de la neuroética

La doble perspectiva metodológica de la neuroética, en tanto que su ejercicio combina ciencia y filosofía, es afín a la propuesta de Korsgaard. Las neurociencias son incapaces de decir nada sobre cómo y por qué justificar nuestras ideas, valores y normas morales, para ello se precisa la autorreflexión, es necesario considerar qué razones existen para justificar moralmente tales ideas, valores y normas. Una de las tareas de la filosofía moral en la neuroética, por tanto, consistirá en llevar a cabo una reflexión que, tomando en cuenta los datos que ofrecen las neurociencias, permita dirimir qué códigos, valores y

normas merecen asentimiento. Además, el carácter público del quehacer filosófico y científico hará que la autonomía de la que habla Korsgaard no sea privada, sino pública, se evitan así los problemas asociados a un uso monológico de la razón, siendo ejercida dialógicamente.

La propuesta de Korsgaard introduce elementos necesarios para nuestra investigación. Entre ellos encontramos la distinción entre los ámbitos explicativo y justificativo de lo moral, una crítica a las posturas de realismo naturalista, que incurren en reduccionismo, y una propuesta sobre el origen de la normatividad, que cristaliza en un procedimentalismo dialógico heredero del pensamiento kantiano. Todos estos elementos subrayan la necesidad de la filosofía moral en el ámbito de la neuroética, apostando por las propuestas procedimentales y dialógicas. Todo ello sienta las bases para una reflexión más profunda sobre las condiciones en las que la fundamentación ética se lleva a cabo desde la filosofía moral. Para ello, analizaremos diferentes propuestas de la ética dialógica, tradición en la cual se inscribe nuestra investigación. En los siguientes capítulos atenderemos a tres figuras que son clave al respecto: Apel y Habermas por un lado, ya que son dos de los máximos exponentes de esta corriente, y Cortina por otro, quien articula una ética de la razón cordial que permite actualizar las propuestas de corte racionalista de sus predecesores, atendiendo a los factores emocionales y estimativos implicados en el diálogo. Todo ello permitirá perfilar los aspectos metodológicos recomendables para articular una neuroética crítica en clave cordial.

Capítulo 8. Un marco interdisciplinar desde la perspectiva dialógica

La relación entre ciencia y filosofía moral, desde una perspectiva dialógica, es abordada de forma profunda por diferentes autores en esta corriente de pensamiento. En el presente capítulo nos propones abordar algunas de las aportaciones de los autores canónicos de esta escuela, Apel y Habermas. El procedimentalismo dialógico que plantean ofrece no solo una teoría ética, sino que ofrece las herramientas conceptuales necesarias para comprender tanto el discurso teórico de la ciencia como el discurso normativo de la ética en tanto que se producen en contextos comunicativos. De este modo es posible englobar ambos bajo un marco filosófico que posibilita compaginar la ciencia y la ética. Este marco es necesario para articular una propuesta metodológica para la neuroética, en la medida en que esta rama del conocimiento se constituye de modo marcadamente interdisciplinar. Gracias al análisis que llevaremos a cabo a lo largo del capítulo podremos ofrecer una comprensión de la ciencia y la ética desde la perspectiva dialógica y comunicativa, especificando la relación que guardan entre ellas, sus competencias, su objeto de estudio, así como una serie de recomendaciones metodológicas que son de gran interés para los objetivos de nuestra investigación.

8.1 Ciencia, ética y el *a priori* de la comunidad de comunicación

Entre ellos, la aproximación de Apel a esta cuestión plantea un enfoque lúcido y valioso que arroja luz sobre algunos puntos que en capítulos anteriores no hemos resuelto. Profundizando en la última conferencia del segundo volumen de la *Transformación de la Filosofía* (Apel, 1985), titulada “El *a priori* de la comunidad de comunicación y los fundamentos de la ética”, tendremos ocasión de ver cómo Apel distingue el ámbito explicativo del justificativo en relación a lo moral, enmarcando ambos en el contexto de una comunidad de comunicación.

El problema de la fundamentación en la relación ciencia-ética occidental.

Apel nos advierte de una paradójica situación que en las sociedades actuales encontramos respecto al problema de la fundamentación. Las consecuencias tecnológicas de la ciencia unifican la civilización a nivel planetario, en esta era científica se hace necesario resolver el problema de la fundamentación ética desde una reflexión racional, de modo que se pueda responder universalmente a problemas que también lo son. Sin

embargo, esa misma ciencia que unifica, por sus consecuencias, a la civilización, también entorpece la resolución al problema de la fundamentación. Advierte Apel que la ciencia prejuzga la idea de validez intersubjetiva, “por la idea cientificista de una «objetividad» normativamente neutral o no-valorativa” (Apel, 1985: 342). Así el reduccionismo científico se configura, como hemos visto ya, como una amenaza para la filosofía moral.

La moralidad endogámica para la que nos dota la naturaleza queda ya muy lejos de las necesidades sociales que en la actualidad encontramos (Ayala y Cela Conde, 2001; Cela Conde, 2001; Waal 1997, 2007; Tomasello y Vaish, 2013; Kohlberg, 1992; Piaget 1987; Cortina, 2012a, 2017). Advertimos, pues, una asincronía entre la expansión de las posibilidades y las consecuencias científico-técnicas, por un lado, y las morales de grupos, por otro, que parecen ser dominantes en comparación con las morales universalistas. Apel distingue tres niveles respecto a los cuales se produce una regulación moral: una microsfera (pareja, familia, grupo, vecindad, etc.), una mesosfera (plano de la política nacional) y una macrosfera (destino de toda la humanidad). Advierte Apel que actualmente la mayoría de los códigos morales se sitúan en los dos primeros niveles, lo cual dificulta la posibilidad de afrontar los problemas globales que la era de la ciencia plantea. La era de la ciencia sitúa a la humanidad frente a un desafío ético común, que debería servir como llamada a una ética de la responsabilidad, acompañada de una validez intersubjetiva de normas.

Hay un segundo aspecto en que la situación es paradójica, esta vez en relación al problema metateórico de la relación entre ciencia y ética. En este caso el problema se deriva de la extendida convicción de que la validez intersubjetiva de argumentos comparte sus límites con la objetividad científica del ámbito de las ciencias reales lógico-matemáticas. Dado que a partir de hechos no podemos deducir deberes morales, y tampoco podemos hacerlo mediante argumentos formales lógico-matemáticos, toda pretensión de validez objetiva queda relegada al ámbito no vinculante de la subjetividad. La fundamentación ética desde las ciencias empíricas queda, de este modo, ilustrada como un imposible. Todo acercamiento de las ciencias a lo moral se producirá, por tanto, en términos descriptivos y explicativos, pero nunca normativos (Korsgaard, 2000; Cortina, 2012a; Evers, 2010; Habermas, 2006; García-Marzá, 2016). Por su parte, existe la acusación de que la filosofía moral tampoco podría resolver el problema de la fundamentación. Una filosofía solo puede ser consecuente cuando abandona la tarea de

la ética, y se ciñe, a lo sumo, a las directrices que marcó la metaética analítica (Albert, 1978), entendida como descripción teórico-científica no valorativa del uso del lenguaje o las reglas lógicas del denominado discurso moral. Toda ética que pretenda defender la posibilidad de una fundamentación es así acusada de dogmática, ideológica y falaz (Apel, 1985: 344-345).

De aquí se deriva que el ser humano ha de superar su historia natural. Este proceso, dice Apel, puede producirse dialécticamente en el proceso histórico. Apel recurre a la propuesta marxista para ilustrar la necesidad de una sociedad hermanada, donde los intereses de grupo, o de clase, hayan sido superados. Sin adentrarnos en más consideraciones respecto a la propuesta marxista, cabe destacar cómo la distinción entre hechos y deberes se diluye al comprender la relación de estos desde una perspectiva dialéctica. En la medida en que no coincidan, los hechos que *son* se oponen a los hechos que *deben ser*, según el agente o la comunidad de agentes morales considere, esto produce una contradicción que puede ser resuelta dialécticamente a través de la acción del agente, al que queda encomendada dicha misión. Así, si el agente obra en atención al deber, podrá procurar que en el futuro los hechos que *son* y los que *deben ser* coincidan, produciendo así una síntesis que supere la contradicción dialéctica, pudiendo de este modo mediar entre teoría y praxis en la situación histórica. El presente y el pasado puede ser abordado desde la experiencia, pero no el futuro, en relación al cual el agente ha de presuponer principios de la acción, no derivados de la experiencia. (Apel, 1985: 346-348; Winch, 2003).

Apel describe un sistema de complementariedad integrado, por un lado, por el objetivismo no valorativo de la ciencia y, por otro, por el subjetivismo de los actos de fe y de las decisiones morales. Esta división puede verse materializada en la separación liberal entre los ámbitos de lo público y lo privado. La propuesta de Cortina (2007), después defendida también por Rawls (1995, 2013), plantea la esfera pública como un ámbito en el que es posible compartir unos mínimos de justicia, mientras que en la esfera privada los máximos de felicidad serán subjetivos. Mientras la esfera privada de máximos no intervenga en la esfera pública de mínimos compartidos una sociedad pluralista será posible y viable. Si bien Cortina y Rawls consideran que los mínimos de justicia presentan por sí mismos un carácter moral, Apel no hace esta distinción al referir a las esferas de lo público y lo privado. Afirma que, en todas las esferas de la vida pública en la sociedad

industrial occidental, la justificación moral recurre siempre a argumentos pragmáticos articulados sobre reglas científico-tecnológicas objetivables. De este modo puede objetivarse una parte de la praxis bajo estándares científico-tecnológicos, con la pretensión de establecer reglas objetivas, con la forma “si-entonces”. (Apel, 1985: 352)

El problema viene dado por el hecho de que, si bien esta objetivación pragmatista de la praxis ofrece los medios para alcanzar unas metas, no puede decir nada sobre la deseabilidad moral de dichas metas (Apel, 1985: 353). La exigencia de poner metas racionalmente se hace aquí patente, quedando revelados los límites y la insuficiencia del concepto científico-tecnológico de una racionalidad que objetiva sin valorar. La mediación científico-tecnológica y no valorativa entre teoría y praxis, en la forma de pragmatismo, no puede asumir la responsabilidad moral que se plantea como exigencia, a lo sumo puede presuponer sus criterios éticos. Si estos criterios, en definitiva, solo pueden producirse en la esfera privada y no vinculante de la subjetividad, encontramos un serio problema. Las convenciones, dice Apel, parecen permitir una traslación de las consideraciones subjetivas al ámbito público, pudiendo servir así de base para una política y un derecho que, por definición, son compartidos. Las decisiones logradas convencionalmente constituyen, así, el fundamento de toda norma intersubjetivamente obligatoria. (Apel, 1985: 353-355; Korsgaard, 2000; Habermas, 2010, 2006; Cortina, 2007, 2009; García-Marzá, 1992)

De aquí se deriva un segundo problema. Esta respuesta, derivada de los presupuestos del sistema de complementariedad, parece resolver el problema, y lo hace de modo que una fundamentación filosófica de una ética universalmente válida parece innecesaria. De este modo el problema que nos atañe, el de la fundamentación, queda disimulado, no resuelto. Toda convención fáctica, si bien resulta valiosa como mecanismo de libertad democrática y de convivencia social, no puede ser entendida más que en sentido contractualista hobbesiano, es decir, como una manifestación racional-teleológica de la prudencia de los individuos, no como un criterio de fundamentación última, por lo cual podría ser efectiva en el mejor de los casos, pero no normativamente obligatoria en sentido moral (Kohlberg, 1992). Debemos insistir en la pregunta acerca de si una fundamentación ética es posible y, de serlo, bajo qué condiciones. (Apel, 1985: 356)

Para que la responsabilidad moral sobrepase la esfera de lo privado es preciso que exista un principio ético que pueda ser considerado intersubjetivamente y que sea

normativamente obligatorio. No queremos decir con ello que las convenciones carezcan de obligatoriedad moral, sino que esta no puede descansar únicamente en la coacción deducida del contrato, descansará más bien en criterios que más allá del pacto, y no necesariamente explicitados, resulten normativamente relevantes para los participantes en el mismo, quienes sentirán la obligación moral de respetar el pacto o las normas establecidas mediante el mismo, incluso cuando la coacción de este no esté operando. (Apel, 1985: 357-358; Cortina, 2010: 113; Cortina, 2009: 43)

Puede observarse, dice Apel, una contraposición entre las aporías ideológicas occidental y oriental. Respecto a Occidente, se postula que todos los individuos son capaces de decisión moral de forma privada y subjetiva, sin embargo resulta imposible una fundamentación intersubjetiva de normas éticas y una solidaridad moral derivada de esta. Respecto a Oriente, se postula una solidaridad de la responsabilidad moral de la sociedad, sin embargo esta no puede estar mediada por las decisiones individuales de conciencia, que resultan superfluas y quedan relegadas a la esfera privada, de forma similar a como sucede en el sistema de complementariedad occidental.

Análisis de las vías posibles para una fundamentación ética.

Existen algunos presupuestos fundamentales de la filosofía analítica que impiden realizar una fundamentación filosófica de una ética normativa universalmente válida. Estos se articulan en tres premisas que mostraremos a continuación:

P1. No pueden derivarse normas, enunciados prescriptivos o juicios de valor a partir de hechos o enunciados descriptivos. Apel referirá a ello como principio o distinción de Hume. (Hume 1988, 1953; Moore, 1983; Cortina, 2012a; Korsgaard, 2000)

P2. La ciencia versa sobre hechos, por ello es imposible una fundamentación científica de una ética normativa. (Gould, 1999; Korsgaard, 2000; Habermas, 2006; Cortina, 2007, 2012a; Salles, 2013, 2016, 2018; Evers, 2011; Nagel, 1985; Cartwright, 2010; Teehan y diCarlo 2004; García-Marzá, 2016)

P3. Solamente la ciencia es capaz de proporcionar un saber objetivo, identificando la objetividad con la validez intersubjetiva. Por ello una fundamentación ética normativa intersubjetivamente válida es imposible.

Para defender la posibilidad de una fundamentación ética debemos, por tanto, cuestionar al menos una de las premisas anteriores. Dos estrategias parecen factibles, cuestionar la primera premisa y cuestionar la tercera premisa.

P1 es cuestionable en la medida en que no todas las ciencias dotadas de contenido empírico son ciencias puras de hechos moralmente no valorativos, quizás lo sean las ciencias de la naturaleza, pero no las ciencias humanas ni la metaética filosófica, que precisarían cierta valoración moral para constituir su objeto de estudio. En el mejor de los casos, solo para las ciencias de la naturaleza podría existir la relación sujeto-objeto, en la que el mundo se da como un conjunto de hechos no valorados. Las ciencias humanas, dice Apel, no han podido hacer efectiva una renuncia a la comprensión valorativa. A pesar de los esfuerzos, a veces fructíferos, de llevar a cabo una comprensión experimental y tecnológica de los objetos humanos, todo intento de una autocomprensión de la praxis humana precisa, para mantener su sentido, la contemplación de unas características valorativas que son constitutivas del objeto.

Las ciencias humanas no pueden desprenderse de la perspectiva valorativa llamada precientífica. Aún los esfuerzos por describir el objeto de forma objetiva conllevan una aproximación apreciativa del investigador que no puede nunca en su interpretación prescindir del todo de elementos valorativos. Más bien su labor consiste en poner en cuestión una aproximación acrítica al objeto, de tal modo que pueda ser conocedor de los aspectos valorativos de su interpretación, depurándolos y aportando además una explicación causal objetiva en la medida que sea posible. Pero esto no garantiza depurar del todo la dimensión valorativa de la aproximación del investigador. Es más, esta aproximación sigue siendo valorativa en algún sentido, es una nueva valoración críticamente mediada. La aportación que se obtiene es el aspecto crítico (Lakatos, 2011). En el caso de la metaética, a pesar de los esfuerzos por desprenderse de los aspectos normativos de una ética filosófica, esta sigue sujeta a prescripciones tales como qué debemos considerar como acción moral o incluso como acción moralmente buena. De este modo, los metaéticos mantienen un componente valorativo del que tampoco pueden desprenderse. A lo cual se suma el problema antes mencionado de que, en todo caso, una aproximación a su objeto, las expresiones humanas en el contexto pragmático, presenta una comprensión que es inevitablemente valorativa (Lenk, 1988, 1982; Apel, 1985: 364). Dicho de otro modo, las ciencias humanas, entre ellas la

metaética, presentan una aproximación hermenéutica a su objeto, en función de la cual una aproximación no valorativa al mismo no es posible (Gadamer, 1991; Conill, 2006). La constitución del objeto se producirá en un compromiso que puede darse comunicativamente y que es normativo tanto a nivel metodológico como moral.

A favor de una hermenéutica normativamente no neutral, Apel afirma que la comprensión de las acciones humanas precisa un compromiso comunicativo, que permita asumir la responsabilidad común por las intenciones de la acción. Además, los datos relevantes solo se constituirán como tales a la luz de teorías, que no serán neutrales normativa ni metódicamente. Así ocurre también con las ciencias empíricas. En el caso de las ciencias humanas estos datos siempre vendrán constituidos desde un enfoque hermenéutico, que será comunicativo y autorreflexivo. Con la única reserva de que es posible un ulterior distanciamiento y neutralización. (Apel, 1985: 365-366)

Estas consideraciones nos permiten afirmar que la distinción de Hume, referida al salto deductivo entre hechos puros y normas, resulta epistemológicamente irrelevante — no falsa— en las ciencias humanas hermenéutico-normativas. Ahora bien, esto no es suficiente para resolver el problema de la fundamentación ética. Apel aporta las siguientes razones que sirven para criticar esta pretensión:

1. La hermenéutica debe presuponer siempre una fundamentación normativa de su comprensión éticamente valorativa.
2. Bajo el supuesto de que la hermenéutica contase con una fundamentación ética, su método no basta a la hora de realizar la aplicación hermenéutica. Para la hermenéutica, en función de su metodología, la justificación de sistemas morales precisa colocarse en las formas de vida socioculturales que queremos comprender, debiendo intentarse una reconstrucción de tales contextos entendidos como situaciones subjetivas que exigen una respuesta. Desde esta perspectiva es metodológicamente imposible excederse más allá de las condiciones de estas formas de vida para ofrecer una justificación ética que pueda ser universalmente válida, lo que permitiría resolver el problema de la fundamentación sin caer en un relativismo moral. A todo ello se sumaría la asincronía derivada de comprender hermenéuticamente contextos y tradiciones morales de grupos o culturas, que no serían compatibles con la necesidad de una fundamentación universal derivada de la unificación cultural que ha traído la globalización de la ciencia, la tecnología y

sus consecuencias. Así pues, la pretensión de universalidad de la hermenéutica, en su aspecto metodológico, ha de ser rechazada como válida para una fundamentación ética universalmente válida. (Apel, 1985: 370)

Sí mantendremos, no obstante, la pretensión hermenéutica de universalidad en su aspecto cuasi-trascendental, en relación al cual el mundo de la vida está siempre lingüísticamente interpretado. El *a priori* de este acuerdo será la condición irrebalsable de posibilidad y de validez intersubjetiva de cualquier construcción teórica, ya sea científica, filosófica o referente al propio lenguaje. Es irrebalsable en la medida en que solo en él puede realizarse el ideal normativo del acuerdo, motivo por el cual ha de ser anticipado en todo caso.

Estas consideraciones permiten concluir, por una parte, que la disyuntiva entre la objetividad no valorativa de las ciencias empíricas y la subjetividad de la moral privada no es sostenible. Por otra parte, podemos concluir también que una fundamentación ética es necesaria antes de que se constituyan las ciencias humanas como *organon* de la ética, ya que, en realidad, en ellas se ha presupuesto ya una fundamentación ética.

Por otra parte, P2 es cuestionable en la medida en que podemos plantearnos si la propia ciencia, pretendidamente objetiva y valorativamente neutral, presupone de forma inevitable la validez intersubjetiva de normas morales. Dicho de otro modo, criticaremos la pretensión científicista de que toda validez intersubjetiva es reducible a la validez objetiva de los enunciados no valorativos de la ciencia. Esta cuestión entronca con las críticas al reduccionismo naturalista que hemos hecho a lo largo de la investigación (Hume 1988, 1953; Moore, 1983; Cortina, 2007, 2012a, 2012b; García-Marzá, 20016; Gould, 1999; Korsgaard, 2000; Habermas, 2006; Salles, 2013, 2016, 2018; Evers, 2010; Conill, 2012).

El método en las ciencias precisa un distanciamiento del universo humano, que permita situar al sujeto enfrentado al objeto. Este distanciamiento es preciso, más aún si cabe, en el caso de las ciencias humanas a la hora de justificar la pretensión de validez respecto a sus proposiciones, incluyendo el caso de las proposiciones éticas. Esta analogía del distanciamiento teórico permite una puesta entre paréntesis de la validez de tales proposiciones, siendo el presupuesto necesario que posibilitará someter estas a crítica, pudiendo así ser discutida y analizada, pretendidamente, desde fuera. El problema aquí lo

encontramos en el hecho de que no podemos desprendernos de “la *exigencia fáctica* de la *pretensión* normativa de validez” (Apel, 1985: 373). La validez problematizada de las normas habrá de reducirse a título de prueba (Cortina, 2007; Habermas, 2010, 2006). Con ello logramos radicalizar la exigencia de una fundamentación ética no dogmática. Aquí reside, análogamente a lo que sucede en la ciencia, la necesidad de un término metaético que permita reflexionar sobre la validez normativa de una proposición, de igual modo que en las ciencias se reflexiona sobre la pretensión de verdad (Habermas, 2006).

Precisamente, el reconocimiento, mediado reflexivamente, de la diferencia entre hechos y normas, obliga a la metaética a considerar la justificación de las exigencias de deber como un problema de «conocimiento» reflexivo y, por tanto, teórico; del mismo modo que ocurre a la hora de justificar las pretensiones de verdad —que son, sin duda, de otro tipo— de las proposiciones teóricas. (Apel, 1985: 373)

Estas consideraciones son de suma importancia para nuestra investigación, la analogía entre el discurso científico y el normativo, distinguiendo la especificidad de cada uno de ellos, será la clave para poder diseñar una neuroética crítica no reduccionista. Habermas también alude a la estructura análoga que existe entre el método de las ciencias naturales y el del discurso normativo, en tanto que se producen al amparo de una comunidad de hablantes que otorgan asentimiento a los enunciados mediante el consenso (Habermas, 2010, 2006). Se hace así necesario recurrir a esa posibilidad constitutiva del pensamiento humano, consistente en situarse en una posición excéntrica, como distanciamiento del mundo y como autodistanciamiento, desde la que parece posible una reflexión del sujeto frente al objeto, necesaria para reflexionar sobre la verdad y sobre la validez normativa de cualquier proposición.

Existe un juego lingüístico trascendental, presupuesto en el discurso teórico, en el cual se inserta el distanciamiento del que hablamos. Este distanciamiento pretende alcanzar un conocimiento de las condiciones de posibilidad del discurso teórico, que no pueden ser suprimidas. Por ello es reflexivo, y no podrá ser dogmático. Un uso dogmático de este distanciamiento produce posturas como el solipsismo metodológico, que queda así refutado, ya que el propio distanciamiento del mundo tiene como condición de posibilidad la pertenencia del sujeto a ese mundo respecto al cual se distancia. La existencia del mundo externo, por ello, no puede ser cuestionada sin anular las condiciones de posibilidad que permiten si quiera formular esa pregunta teórica (Habermas, 2006: 42-43; Apel, 1985: 374). Del mismo modo hay cosas que no pueden

ser puestas en duda en el ámbito de la ética, sin anular las condiciones mismas de la ética. Atendiendo a las condiciones de posibilidad de esta encontramos que “el presupuesto de la validez de normas morales en general es una condición «paradigmática» de posibilidad del juego lingüístico que pretende justificar normas” (Apel, 1985: 375). Así, casos como el de la tesis metaética de la neutralidad quedan refutados, del mismo modo que sucede con el solipsismo metodológico.

Llegados a este punto, podemos dar un paso más. Una vez aceptado el punto de vista metodológico de la reflexión trascendental, podemos preguntarnos no solo por las condiciones de posibilidad y validez del discurso normativo, sino también del científico, pretendidamente objetivo y no valorativo. El adecuado análisis de esta cuestión permitirá en primer lugar caracterizar metodológicamente ambos discursos y sus competencias. En segundo lugar, permitirá refutar la tesis de que existe una imposibilidad de validez intersubjetiva para las normas morales. Y por último permitirá responder a la pregunta de si es posible determinar las posibilidades de validez y objetividad no valorativa de las ciencias naturales sin que estas presupongan ya la validez intersubjetiva de normas morales. (Apel, 1985: 376)

El carácter normativo de las ciencias naturales.

Retomando las cuestiones que cerraban el apartado anterior, a continuación trataremos de defender la tesis de que las ciencias naturales, pretendidamente no-valorativas, presentan la presuposición de la validez intersubjetiva de normas como condición de posibilidad. A partir de ello podremos refutar la tesis de la subjetividad de las normas morales y podremos distinguir y especificar la constitución metodológica de los discursos científico y normativo. Que será necesaria para elaborar una propuesta metodológica adecuada para una neuroética crítica.

Si logramos demostrar que la validez intersubjetiva de normas morales es una condición de posibilidad y validez propia de las ciencias naturales, varias consecuencias se derivarían de ello. En primer lugar, encontraríamos que toda pretensión de objetividad no valorativa en las ciencias naturales descansa, en último término, en la aceptación intersubjetiva de normas. Esto, a su vez, permitirá refutar el argumento según el cual la posibilidad de una ética descansa en que su forma de validez sea reducible a la objetividad no valorativa de las ciencias empíricas.

En tal caso la ciencia natural presupondrá siempre un acuerdo metodológico en relación a la dimensión de la intersubjetividad, de carácter trascendental. Según esto, antes de que el discurso científico tenga lugar se han de dar unas condiciones de posibilidad del mismo, que serán de carácter trascendental, en el que la validez intersubjetiva de normas se da como presupuesto, y sin el cual la ciencia no puede producirse con sentido. Estas normas presupuestas no suponen meros imperativos hipotéticos, sino que proporcionan una respuesta no solo a *cómo* debe hacerse la ciencia, sino a si la ciencia *debe ser, y por qué*. Cabe añadir que la presuposición de la validez intersubjetiva de normas, si bien se produce en la ciencia, no es exclusiva de esta, estando presente, en realidad, en toda discusión acerca de problemas (Apel, 1985: 377).

La ética como presupuesto de la lógica.

La lógica, especialmente en el caso de la ciencia, ha de concebirse como una ciencia normativa (Apel, 1985), lo cual supone que las ciencias naturales no se constituyen como unas ciencias empírico-analíticas no valorativas, sino más bien como una ciencia normativo-hermenéutica, siendo la lógica uno de sus principios de valoración normativa, que constituyen a su vez una de sus condiciones de posibilidad. No obstante, no ha de tomarse el principio normativo de la lógica como un principio suficiente para la fundamentación ética. La lógica, entendida como teoría normativa del correcto uso del entendimiento, puede ser usada instrumentalmente, sin perseguir por ello, y tampoco fundamentar, una ética. Dicho de otro modo, la lógica no implica una ética, pero sí que la presupone como condición de posibilidad, como sucede con toda ciencia.

Esta tesis está apoyada por el hecho de que no se puede comprobar la validez lógica de argumentos sin presuponer una comunidad de hablantes o de pensadores con capacidad para el acuerdo intersubjetivo a través del consenso. Incluso en el ejercicio monológico de un pensador se está interiorizando la estructura del diálogo, esta vez con uno mismo. Es decir, se está interiorizando el diálogo de una potencial comunidad de comunicación. Todo pensamiento monológico debe su validez, por tanto, a la posible justificación de las proposiciones en el marco de una comunidad de argumentación (Apel, 1985; Cortina, 2007; Habermas, 2010; García-Marzá, 1992). Así pues, la justificación, incluso monológicamente producida, presenta un presupuesto normativo, en función del cual la mentira o la no consideración crítica de cualquier argumento no pueden ser defendidas. Apel concluye, al respecto, que “en la comunidad de argumentación se

presupone que todos los miembros se reconocen recíprocamente como interlocutores con los mismos derechos” (Apel, 1985:380). Estos presupuestos normativos se constituyen, así, como condición de posibilidad de la lógica, de la ciencia y de la ética, en tanto que discursos sujetos a unas condiciones de argumentación que no pueden ser negadas sin perder con ello el sentido.

Toda expresión lingüística y toda acción con sentido son virtualmente un argumento. Por ello, todos los seres capaces de comunicación lingüística y de argumentación han de ser reconocidos como tales, argumenten o no fácticamente, de este reconocimiento dependerá no socavar la condición de posibilidad antes descrita, pues todas las aportaciones virtuales al diálogo han de ser consideradas. Por tanto, cada enunciado sobre hechos, como ocurre en la ciencia, en la medida en que ha de ser lógicamente justificado, ha de reconocer a su vez este presupuesto comunicativo, que es normativo. Así el enunciado, performativamente, al ser afirmado, reconoce este presupuesto en la medida en que lo presupone necesariamente, presuposición que implica una serie de aspectos éticos de forma insoslayable.

Por todo ello, no pueden justificarse argumentos, ni si quiera en el nivel monológico, prescindiendo totalmente del nivel dialógico. Todo argumento, científico, lógico o ético, queda de este modo ligado necesariamente a la dimensión pragmática. Los argumentos habrán de ser entendido en todo caso como pretensiones de sentido y validez que habrán de ser explicitadas y validadas en un contexto dialógico, al menos virtualmente. Alcanzamos con ello una consideración pragmático-trascendental de las condiciones de posibilidad de la lógica y de la ciencia, que remite directamente a la comunidad de comunicación como un *a priori* irrebable de ambas.

La ética de la argumentación presupuesta tanto por la lógica como por la ciencia puede responder a la pregunta sobre si estas deben de ser y por qué. Esta ética no es solo una condición de posibilidad para la justificación o validación de las proposiciones de las ciencias o de la lógica, sino que se constituye como una exigencia hacia todos los miembros de una comunidad comunicativa, desde cada uno de ellos en relación con los demás, configurándose así como una obligación moral. Si la justificación de normas no queda configurada como deber, la condición de posibilidad que hemos descrito como necesaria se diluye, y con ello, consiguientemente, toda posibilidad de una ciencia o una lógica con sentido. Así pues, la condición de posibilidad de la ciencia y de la lógica

implica que esta condición ha de ser aceptada siempre y cuando queramos hacer ciencia y lógica con sentido.

El *a priori* de la argumentación.

Un posible problema para la fundamentación racional de la ética sería el siguiente. Si toda lógica presupone ya una ética, y toda fundamentación racional de la ética presupone ya la validez de la lógica, esto podría generar un regreso al infinito o un círculo en la argumentación que imposibilitarían la propia fundamentación última de la ética (Apel, 1985; Popper, 1962; Albert, 1978; Cortina, 2007; Korsgaard, 2000). Ahora bien, este problema puede resolverse si la fundamentación que perseguimos no es entendida en el marco de un sistema axiomático, sino en el de una reflexión trascendental sobre las condiciones de posibilidad y validez de toda argumentación. Apel denuncia la poca atención que en la discusión ha recibido la reflexión trascendental como método específicamente filosófico, ocupando un primer plano la perspectiva analítica que conduce a pensar que el problema de la fundamentación es un problema de presupuestos lógicos, sintáctico-semánticos, para las aserciones o proposiciones. El problema radica en que la competencia reflexiva no puede autocontemplarse en términos de lenguajes formalizables, toda autorreferencia supone una contradicción (Russell, 1966). Es precisamente en esta imposibilidad lógico-semántica de objetivar las condiciones de posibilidad subjetivas de la argumentación donde se hace patente la necesidad del saber autorreflexivo en torno al sujeto pragmático-trascendental de la argumentación, desde el cual es posible una teoría de la prueba que evite de algún modo el regreso al infinito que impediría la fundamentación (Cortina, 2007).

Hemos de considerar, pues, como una vía para la fundamentación ética última, la posibilidad de una pragmática trascendental del lenguaje, que permita al sujeto reflexionar sobre las condiciones de posibilidad y validez de la argumentación, presupuestas en toda situación de habla, incluyendo el pensamiento monológico, como situación de habla interiorizada. Al reflexionar sobre las condiciones de posibilidad de la comprobación crítica, afirma Apel, encontramos algo irrebasable en sentido pragmático-trascendental. La tarea propia de la filosofía consiste, según Apel, en reconstruir las condiciones necesarias de la argumentación humana, tan completamente como sea posible. La perspectiva argumentativa de Apel, que es en todo caso dialógica, se diferencia de eso mismo respecto a la propuesta kantiana, que en virtud de un solipsismo

metódico cree posible el análisis trascendental desde una perspectiva monológica (Kant, 1992, 2000). Para Apel la unidad intersubjetiva de interpretación determina el sentido y el consenso respecto a la verdad y respecto a la validez, cada una de ellas en referencia al discurso que le corresponde, a saber, el científico y el normativo, respectivamente (Cortina; Habermas; Apel). Así pues, la prueba crítica y las condiciones trascendentales partirán de la argumentación como un punto de referencia irrebasable. (Apel, 1985: 391)

Apel contrapone su propuesta al racionalismo crítico de Popper (1962, 1967, 1992). El *a priori* de la comunidad de comunicación guarda ciertas semejanzas con el *criticist frame* popperiano, pero también unas diferencias significativas. Partiendo de un rechazo de la fundamentación última, Popper concluye que no es posible atribuir una prioridad al *criticism frame* frente al irracionalismo, y que en cualquier caso la decisión por una u otra opción resulta arbitraria y basada en un acto de fe (Popper, 1992). Apel, sin embargo, reconoce que la realización práctica de la razón a través de la buena voluntad precisa siempre de un compromiso que no puede ser demostrado, ello no significa que hayamos de reconocer un rechazo a una fundamentación racional del compromiso más primario con la razón. Esa identificación solo podría darse si identificamos la fundamentación con la deducción, ignorando la relevancia de la reflexión, capaz de señalar las condiciones pragmático-trascendentales del sujeto de argumentación. Así pues, cualquiera que entre en el juego de la discusión, solo puede ser consistente semántico-pragmáticamente, como revela la reflexión trascendental, aceptando las condiciones de posibilidad de la discusión misma, que son, entre otras cosas, racionales. La decisión de optar por el irracionalismo ni tan si quiera se produce en el marco de la argumentación, y la postura de quien tome tal decisión resultará, por ese motivo, irrelevante para la discusión. Con ello queda refutada la tesis popperiana sobre el carácter irracional y moral de la decisión, esta será moral no en virtud de su irracionalismo, sino por reforzar voluntariamente los principios, reflexivamente establecidos, de la discusión. Parece, además, que Popper considera posible pensar y decidir con sentido de forma previa al reconocimiento, al menos implícitamente, de las condiciones de la argumentación en una comunidad de comunicación. Ahora bien, dice Apel, apelando al segundo Wittgenstein (1988), que las condiciones, lógicas y morales, de la argumentación crítica y comunicativa son ya presupuestas por el sujeto que pretenda obrar con sentido, o en este caso, pensar y decidir con sentido, aun si lo hace monológicamente en su foro interno.

La norma fundamental.

La voluntad racional ha de presuponerse en cada discusión sobre la fundamentación para que esta esté dotada de sentido. Asimismo, la voluntad de argumentar es igualmente necesaria para preservar el sentido. Además, esta voluntad no está empíricamente condicionada, sino que la reconocemos como una condición de posibilidad trascendental de toda discusión. Si buscamos una fundamentación en sentido incondicional, es decir, sin depender de las condiciones empíricas, la norma fundamental será incondicionada o categórica, y podremos hallarla en la voluntad de argumentación, en la medida en que se da *a priori* y como condición de posibilidad de la discusión. Así pues, la objeción de que nuestra propuesta podría fundamentar solo imperativos hipotéticos no resiste la crítica.

Otra objeción, sostenida en contra de nuestra propuesta, plantea que partiendo de un hecho determinado, el haber aceptado las condiciones de posibilidad y de argumentación, que incluyen normas morales, no puede deducirse una norma moral. Apel reconoce que la aceptación de normas es una condición necesaria, pero no por ello suficiente, para la validez moral de las normas. Si bien en todo contrato, para ser realmente vinculante, se presupone la aceptación de unas normas, la aceptación misma del contrato no servirá en ningún caso como fundamentación última de dichas normas (Cortina; Kohlberg, 1992). Dicho de otro modo, el contrato permite aceptar unas normas en un acuerdo, pero no dice nada sobre la legitimidad moral de dichas normas.

Frente a ello argumentaremos que la aceptación de las reglas de juego de una comunidad de comunicación no es un hecho empírico, del que se pretendan deducir unas normas, violando la distinción de Hume entre hechos y normas (Hume, 1988, 1953). La aceptación de dichas reglas es una condición de posibilidad y validez de toda argumentación, que en el caso del discurso científico permitirá a su vez toda comprobación de hechos empíricos, mientras que en el caso del discurso normativo permitirá la crítica y asentimiento ante la validez de normas morales (Apel; Korsgaard; Cortina; Habermas), sin por ello cometer un salto falaz entre hechos y normas.

Yendo un paso más allá, es precisa la aceptación de una norma moral fundamental, en la medida en que constituya una condición de posibilidad de toda argumentación, así como de toda autocomprensión válida. Es decir, la norma fundamental presupuesta en las

condiciones de posibilidad de toda argumentación posible se configura como un *factum* de la razón en sentido kantiano, y no como un *factum* humeano, en sentido empírico. Es por ello que su carácter es categórico y no hipotético. Si bien el discurso kantiano, como dice Apel, está expuesto a la crítica de Hume, presentando algunos problemas dignos de ser considerados (Apel, 1985: 396-397), Apel propone una superación de la propuesta kantiana interpretándola y reconstruyéndola en sentido dialógico (Apel; Korsgaard; Habermas; Cortina), en el cual el *factum* de la razón antes aludido representaría una implicación del *a priori* de la argumentación, tal y como hemos expuesto. De este modo, Apel otorga cierta ventaja a la propuesta kantiana frente a otras como el reduccionismo empirista o el decisionismo.

Así pues, frente a posturas como las de Lorenzen (1969), Popper (1962, 1967, 1992) y Schwemmer (1980), que creen que no podemos justificar los criterios mediante los que fundamentamos normas morales, considerando que un acto de fe es inevitable, puede argumentarse, junto a Fichte (1910) y Apel (1985), que la propia participación en una comunidad de comunicación permite inteligir tanto el principio moral como su validez, mediante una reconstrucción práctica de la razón, en la medida en que se da como condición de posibilidad de toda argumentación. Todo participante en la discusión, por este motivo, ha aceptado ya tanto este principio moral como su validez, para poder argumentar con sentido. Desde esta perspectiva pragmático-trascendental podemos al mismo tiempo apoyar la validez de una norma fundamental y refutar la afirmación de que un acto de fe es necesario para justificar los criterios que permiten justificar la validez de las normas morales. Por tanto, la exigencia de fundamentar una norma moral básica o principio de la ética no carece de sentido ni se basa en la arbitrariedad de un acto fe, sino que descansa en un reconocimiento de las condiciones pragmático-trascendentales de toda argumentación posible, habiendo realizado así una reconstrucción dialógica de la propuesta kantiana.

El contenido de la norma ética fundamental.

Llegados a este punto, podemos preguntarnos qué logros permiten extraer las consideraciones anteriores en relación con una posible fundamentación ética válida universalmente. Apel atiende aquí a Pierce (1965), quien reconoce en la comunidad de científicos de la naturaleza un compromiso moral y una aceptación de las condiciones de posibilidad trascendentales de la comunidad de comunicación, en este caso en relación a

la búsqueda de la verdad, que en el caso del discurso ético correspondería al establecimiento de la validez normativa. Apel, junto a Pierce, indica que el científico ha de reconocerse como miembro de una comunidad de comunicación virtualmente ilimitada, correspondiente a la comunidad científica, siendo cada científico conocedor de que la verdad no será alcanzada por los individuos que forman la comunidad de comunicación, sino que es un objetivo de dicha comunidad, al margen de si este se realiza o no en la vida del individuo concreto.

Esta necesaria adscripción del científico a la comunidad de comunicación muestra que la comunidad científica presupone la comunidad de argumentación. Hay que matizar a este respecto que esto no significa que se identifique con ella, puesto que no toda comunidad de comunicación es científica. La comunidad científica presupone y se configura como comunidad de comunicación, sin que por ello signifique que no existan otras comunidades de comunicación que no sean científicas. Existen, pues, comunidades de comunicación no científicas, siendo una de ellas la correspondiente al discurso moral. Si bien los objetivos de las diferentes comunidades de comunicación pueden diferir entre unas y otras, coinciden las condiciones de posibilidad pragmático-transcendentales de las mismas, en la medida en que la comunidad de argumentación es un presupuesto necesario de todas ellas.

Esto hace que los discursos propios de cada comunidad de comunicación presenten una especificidad propia respecto a objetivos, competencias y metodología, si bien todas ellas comparten las condiciones de posibilidad en tanto que comunidades de comunicación. Mientras que la ciencia se ocupa de la búsqueda de la verdad, la ética se ocupa de la búsqueda de la validez normativa, o más concretamente, de los criterios que permitan justificar la validez normativa, aunque su vigencia sea siempre revisable (Apel; Cortina; Habermas). Cabe señalar aquí que también la aceptación y vigencia de las teorías y las hipótesis científicas, análogamente, es siempre revisable. Así pues, el *a priori* de la comunidad de argumentación contiene la exigencia de justificar todas las afirmaciones que se hagan, esta exigencia se satisfará desde unos criterios de racionalidad y reconocimiento mutuo entre los interlocutores, sean estas afirmaciones científicas, relativas a la verdad, o éticas, relativas a la validez normativa.

Los miembros de la comunidad de comunicación, de este modo, están obligados a reconocer las exigencias de todos los miembros virtuales de dicha comunidad. Esto

supone reconocer todas las necesidades humanas virtuales, entendidas como exigencias que podrían argumentarse interpersonalmente, en virtud de lo cual resultan éticamente relevantes. Así pues, afirma Apel, “todas las *necesidades* de los hombres, que puedan armonizarse con las necesidades de los demás por vía argumentativa, en tanto que *exigencias* virtuales, tienen que ser de la incumbencia de la comunidad de comunicación” (Apel, 1985: 404).

El principio fundamental de una ética de la comunicación queda dibujado en el fragmento anterior. La norma ética básica se presenta como una obligación para todos aquellos que han adquirido competencia comunicativa, y que mediante el consenso intersubjetivo han de formar solidariamente la voluntad en cada asunto que afecte a alguno de los otros miembros de la comunidad (Apel, Habermas, Cortina). Esta norma se da en la esfera pragmático-trascendental relativa a las condiciones de posibilidad de toda discusión, y no queda justificada meramente por la aceptación de un convenio como sucedería en el modelo contractual (Cortina, 2010: 113; Cortina, 2009: 43), o como sucede en los primeros estadios del desarrollo moral (Kohlberg, 1992; Piaget, 1987). Toda decisión, incluso la que se produzca a nivel monológico de forma individual y subjetiva, incluyendo aquí el solipsismo metódico —también en el ámbito de la ética—, quedará así mediada de antemano por el *a priori* de la comunidad de comunicación en el que la intersubjetividad y la virtual explicitación pública de los argumentos son un criterio irrebasable de validez en la formación de la voluntad.

El establecimiento de este principio conlleva la necesidad de desarrollar un *método de la discusión moral*, en palabras de Apel, que sea institucionalizado, además, eficazmente, en términos jurídico-políticos. Aquí se presentan algunas dificultades, si bien la fundamentación de una ética de la comunicación se produce en términos idealizados, su realización se producirá siempre en una situación histórica concreta, en la que todo miembro de la comunidad permanece inevitablemente ligado a una situación personal y social concreta, en la cual sus intereses pueden interferir, y lo harán casi con certeza, en el ideal de la comunidad de argumentación, asumiendo una responsabilidad moral específica, que se alejará siempre del ideal. La consideración ideal del principio moral no atiende, ni puede hacerlo, a la situación moral específica de quienes han de tomar decisiones en un contexto, una situación y unos plazos determinados, sujetos a la necesidad de contemplar las consecuencias posibles o probables en tal contexto.

Contemplar los aspectos cordiales implicados en la argumentación ayudará a reducir las dificultades que este problema supone, aunque no logre resolverlo del todo (Cortina, 2007). Cabe preguntarse si al menos, desde la propuesta dialógica, pueden extraerse unos principios regulativos que guíen la decisión en cualquier contexto. Apel trata de responder afirmativamente, bosquejando cómo el *a priori* de la comunidad de comunicación puede guiar, al menos estratégicamente, la acción moral.

El *a priori* de la comunidad de comunicación no es puramente idealista, pero tampoco materialista. En lugar de ello se sitúa como una dialéctica entre ambos. El miembro de la comunidad de comunicación presupone dos cosas al mismo tiempo:

- a) una comunidad real de comunicación, de la que ha pasado a formar parte mediante un proceso de socialización y endoculturación.
- b) una comunidad ideal de comunicación, en cuyo seno todo argumento puede ser tenido en cuenta en condiciones tales que se pueda juzgar su verdad o validez de forma adecuada, en relación con las exigencias del resto de miembros de la comunidad.

La comunidad real no se identifica con la ideal, pero, sin embargo, quien argumenta tiende a presuponer la comunidad ideal en la real, considerando que en la comunidad real existe la posibilidad de una realización de la ideal, aunque sabe que la primera dista mucho de ser como la segunda. El presupuesto trascendental de la comunidad de argumentación contiene así una contradicción dialéctica. No debemos tratar de disolver la contradicción separando los dos supuestos, el ideal y el real, más bien debemos mantenerla como contradicción dialéctica, sirviéndonos de esta para orientar la acción moral. Para ello debemos postular la disolución histórica de la contradicción, aunque no por ello signifique que pensemos que va a producirse. A partir de este postulado, podremos contar con dos principios regulativos fundamentales, que permitan guiar el obrar moral:

- a) hemos de tratar de garantizar la supervivencia de la humanidad como comunidad de comunicación.
- b) debemos intentar realizar la comunidad de comunicación ideal en la real.

El primero de los principios es una condición de posibilidad del segundo, y el segundo dota de sentido al primero. La estrategia de la supervivencia, que se plantea

como un ideal regulativo desde el *a priori* de la comunidad de comunicación, se traduce a largo plazo, dice Apel, como una estrategia de emancipación. De este modo pretende Apel atender a la necesidad de responsabilizarse solidariamente ante la amenaza de las macroconsecuencias derivadas de la expansión mundial de la ciencia y la técnica. En la medida en que el *a priori* de la comunidad de comunicación presupone una ética, aquel quehacer científico que la desatienda estaría socavando las condiciones de posibilidad de las que depende su existencia. Apel, además, indica que en sentido neo-marxista no ortodoxo, la sociedad de clases quedaría superada, emancipatoriamente, en la realización de la comunidad ideal de comunicación. De este modo, como indica Cortina, “sigue teniendo la filosofía la función crítica y liberadora que, por su mismo estatus epistemológico, no pueden ejercer las ciencias” (Cortina, 2010: 31). Dicho de otro modo, se acabaría así con todas las asimetrías que el dialogo interpersonal presentaría en función de las condiciones sociales determinadas. Incluso en el momento en que un individuo ha de tomar la decisión de qué posición moral concreta adopta en su situación histórica, momento arriesgado en el que la decisión no puede fundamentarse, o al menos no totalmente, la mejor opción, dice Apel, será regir siempre la decisión atendiendo, mediante autocomprensión reflexiva, a la crítica posible que tomaría forma en el seno de una comunidad ideal de comunicación.

8.2 Perspectivas metodológicas y epistémicas en clave dialógica

Para continuar y profundizar en los aspectos metodológicos de una neuroética crítica y no reduccionista en clave dialógica es preciso profundizar en la propuesta de Habermas. Este comparte con Apel la perspectiva propia de las éticas dialógicas, aunque ambos difieren entre sí en algunos aspectos que son específicos del pensamiento de cada autor. En lo que concierne a la dimensión metodológica tanto del discurso científico como del normativo, Habermas aporta algunas reflexiones que serán cruciales.

Este compila una serie de conferencias en su obra *Entre naturalismo y religión*. La reflexión sobre las relaciones entre tales elementos conduce a Habermas a tratar, en algunos puntos, el tema del naturalismo en tanto que puede tomar la forma de una pretensión científicista. Si bien Habermas reconoce el sustrato biológico del ser humano y las competencias del naturalismo en tanto que capaz de estudiar ese sustrato para su comprensión, así como el hecho de que todas las operaciones del espíritu —en tanto que dimensión mental subjetiva— dependen de ese sustrato, nos advierte de que hay que ser

cautos en torno a cómo naturalizar el espíritu. Lo que él sugiere es que una adecuada comprensión naturalista de la dimensión cultural humana, asociada al espíritu, ha de dar cuenta de la constitución intersubjetiva así como del carácter normativo de las operaciones de este, regidas por reglas (Habermas, 2006: 9).

Además, junto a la cuestión del naturalismo, Habermas se ocupa también de reflexionar sobre el establecimiento de lo normativo en la esfera pública, mediante reflexión discursiva y dialógica, donde la primera y segunda persona ocupan un papel preponderante, quedando la tercera persona, que presenta una perspectiva de observador, relegada al ámbito de la ciencia descriptiva y explicativa (Cortina, 2007, 2012a; Evers, 2010). Retoma Habermas en esta obra su propuesta de una ética dialógica (Habermas, 2010) y la pone en relación con cuestiones tales como el naturalismo o la dualidad de perspectivas metodológicas, siendo estas la del partícipe y la del observador, que se presenta como ineludible e irreductiblemente necesaria para acercarnos a un conocimiento adecuado del mundo. Dos son, pues, los focos de nuestra reflexión en torno a Habermas, el carácter dialógico y discursivo del establecimiento de lo normativo, así como también de la ciencia, cuestión que ya hemos podido tratar con Apel, y la relación entre un concepto de naturalismo *sui generis*, que llamaré *naturalismo blando*, con una dualidad de perspectivas que nos viene impuesta en la búsqueda de un conocimiento adecuado sobre el mundo.

Ética y ciencia como comunidades de comunicación.

A lo largo de las diferentes conferencias que componen *Entre naturalismo y religión* la cuestión de la ética dialógica o discursiva aparece recurrentemente, y no es de extrañar dado que la propuesta habermasiana se articula en torno a ella. Resulta especialmente interesante el tratamiento de la cuestión que el autor hace en las conferencias 1, 2 y 3, que componen la primera parte de la obra, titulada “La constitución intersubjetiva del espíritu guiado por normas”, en la que recopila los principales aspectos de toda su obra (siendo la referencia más relevante: Habermas, 2010). De estas tres conferencias la primera indaga en las raíces biográficas del pensamiento habermasiano, la segunda se ocupa de las cuestiones de la acción comunicativa y la razón sin trascendencia, la última consiste en un análisis de la propuesta habermasiana puesta en relación con la de Apel. De las tres conferencias atenderemos especialmente a la segunda, en la que se ponen sobre el tapete cuestiones cruciales para acercarnos al pensamiento de

Habermas y dialogar con él. La segunda parte de la obra, “Pluralismo religioso y solidaridad ciudadana”, se adentra en cuestiones relativas a los fundamentos prepolíticos del Estado democrático de derecho y a la religión en la esfera pública, dedica una conferencia a cada uno de estos temas. En la tercera parte el autor se ocupa de la relación entre naturalismo y religión, retomaremos algunas de las conferencias de esta parte más adelante. Por el momento vamos a centrar nuestro análisis en las conferencias de la primera parte.

La primera conferencia, con motivo de rastrear las raíces biográficas del pensamiento de Habermas, expone algunas de las ideas principales de su propuesta. Además de los datos biográficos que en ella se relatan encontramos una exposición de las ideas centrales del autor. La conferencia cumple muy bien su cometido como introducción al libro, además atender a ella también nos servirá para introducirnos en las cuestiones principales a las que nos interesa atender.

Habermas declara que tanto el espacio público como la esfera pública política han sido aspectos centrales de su obra, los cuales él comprende no como un espacio de espectadores y oyentes sino como un espacio de hablantes y receptores, donde tiene lugar un intercambio de razones que persigue una convergencia en creencias e ideas. Junto a la esfera pública, el discurso y la razón son también elementos centrales en toda su obra, pues mediante un discurso racional, que haga públicos los puntos de vista de los interlocutores en un diálogo, podrá llegarse a acuerdos, que serán especialmente relevantes en lo que concierne al establecimiento de lo normativo. Sostiene Habermas que la interioridad del sujeto individual presenta ya un carácter exterior en la medida en que su espíritu recibe tanto la estructura como el contenido de su vinculación con el espíritu objetivo a través de la intersubjetividad que supone su relación con los otros socializados (Ayala y Cela Conde, 2001; Waal, 1997, 2007; Tomasello y Vaish, 2012; Kohlberg, 1992). La externalización del centro interior mediante la comunicación en las relaciones personales es necesaria para el crecimiento vital y personal de forma autoconsciente. Es ese ámbito comunicativo el que define quienes somos en contraste y relación con los demás (Evers, 2010). La subjetividad del otro es condición necesaria para la autocomprensión individual, como algo puesto en relación pero diferenciado de ese otro que al mismo tiempo choca conmigo y me configura como individuo.

Esa interrelación subjetivadora e individualizadora que es a la vez intersubjetiva y comunicativa supone una serie de riesgos asociados a la reciprocidad denegada, Habermas destaca el papel crucial de la moral del igual respeto como medio para atajar los problemas que suponen la marginación de unos individuos respecto a otros. La puesta en público de todos y cada uno de los afectados por las normas, así como el reconocimiento de su igual derecho a la participación en un diálogo que se constituye como procedimiento para establecer normas, pretenden garantizar la disolución de tales problemas.

Cuanto más avanza la individualización hacia dentro tanto más profundamente se enreda el individuo, por así decirlo, hacia fuera en una red cada vez más densa y frágil de relaciones de reconocimiento recíproco. Con ello se expone a los riesgos de la reciprocidad denegada. La moral del igual respeto para todos quiere atajar estos riesgos, pues se define sobre el objetivo de la abolición de la discriminación y de la incorporación de los marginados en la red de deferencia recíproca. Normas de convivencia que pueden instaurar solidaridad también entre extraños dependen de una aprobación general. Tenemos que introducirnos en discursos para desarrollar tales normas, pues los discursos morales dan la palabra por igual a todos los afectados. Los discursos espolean a los participantes a adoptar las perspectivas de todos los demás. (Habermas, 2006: 24-25).

Esta perspectiva es afín en términos generales a la de Korsgaard (2000) y a la de Apel (1985), a las que ya hemos atendido, y a la de Cortina (2007, 2009, 2012a), en la que profundizaremos en el próximo capítulo. El realismo procedimental que Korsgaard y la comunidad de argumentación de Apel sugieren una puesta en público de las razones privadas de cada individuo que pueda empujar al resto de individuos a una reflexión racional que permita asentir ante aquellas ideas, valores y normas que cuentan con buenas razones, rechazando aquellos que no presenten tales razones. Por su parte Cortina, desde la propuesta de una ética de la razón cordial, sugiere tener en cuenta los factores emocionales y estimativos implicados y relevantes para la discusión. Cortina señala que en estos autores la razón kantiana, en su uso teórico y específicamente en su uso práctico, se descentraliza y se desublima, pasando a hablar de una razón situada que se encarna en la práctica comunicativa cotidiana, penetrando en los contextos de acción de la realidad social. Mientras que en Korsgaard esta cuestión es tratada de forma más sucinta, Apel y Habermas, atendiendo este último a McCarthy (1992), la tratan explícitamente. Esta es la cuestión con la que abre su segunda conferencia en *Entre naturalismo y religión*. El sujeto en su dimensión de sujeto cognoscente y también en su dimensión de sujeto normativo se

descentraliza, ahora ya solo es cognoscente y normativo en contextos comunicativos, donde la reflexión y el asentimiento y donde la adopción de creencias y normas se produce intersubjetivamente. Frente a propuestas de liquidación del concepto mismo de razón, como la que esgrime Davidson (2001, 2004), otros autores como Dummet (1993) o Brandom (1994), y por supuesto Korsgaard, Apel, Habermas y Cortina, proponen una desvinculación del concepto de razón de su marco mentalista trasladándolo a un plano intersubjetivo como razón comunicativa, reconstruyendo la normatividad en la práctica del entendimiento (Habermas, 2006: 63).

En este contexto de descentralización de la razón pura kantiana es donde cobra sentido el planteamiento de McCarthy cuando habla de los análogos práctico-sociales de las ideas kantianas de la razón (McCarthy, 1992). Según McCarthy estos son tres presuposiciones pragmático-formales de la acción comunicativa, a saber:

- a) la suposición de un mundo objetivo.
- b) la suposición de mutua racionalidad entre los interlocutores.
- c) la validez incondicionada de los enunciados resultantes mediante los actos dialógicos y comunicativos de habla.

Habermas añade más, y los relaciona con conceptos kantianos (Kant, 1978). La suposición de un mundo objetivo la relaciona con las ideas cosmológicas de la unidad del mundo. La suposición recíproca de la racionalidad de los interlocutores, a la que añade Habermas la responsabilidad, está relacionada con la idea de libertad como postulado de la razón práctica y como suposición pragmática de racionalidad en actores responsables (Kant, 1992, 2000). La incondicionalidad de validez de los enunciados resultantes de los actos comunicativos de habla se relaciona con el movimiento totalizante de la razón. A estas tres suposiciones añade Habermas —si bien se halla también en McCarthy, aunque Habermas no lo dice explícitamente— la razón como capacidad de principios que se relaciona con el discurso racional en tanto que tribunal o “foro último e irrebasable de toda posible jurisdicción” (Habermas, 2006: 36). Encontramos en la razón una condición de posibilidad irrebasable del discurso, siendo a su vez una de las condiciones que constituyen el *a priori* de la comunidad de comunicación al que atendimos en el capítulo anterior (Apel, 1985).

Asegura Habermas que estas presuposiciones idealizantes de la acción comunicativa permiten la comprensión respecto al papel fáctico que juegan los supuestos contrafácticos.

Como suposiciones idealizantes que no podemos evitar cuando entramos en procesos de entendimiento mutuo, son *realmente efectivas*, ya que organizan la comunicación; y, al mismo tiempo, son *contrafácticas*, en tanto que apuntan más allá del límite de las situaciones presentes. Como resultado, las ideas práctico-sociales de la razón son, al mismo tiempo, “inmanentes” y “trascendentes” respecto a las prácticas constitutivas de las formas de vida. (Habermas, 2006: 34)

Aun con todo, reconoce que las ideas de la razón no pueden ser traducidas de forma directa desde el lenguaje filosófico trascendental al de la pragmática formal. En esta cuestión vemos reflejado uno de los objetivos principales de Habermas, que, como Apel (1985), investiga sobre las condiciones pragmático-trascendentales de la argumentación.

Cabe destacar en este punto la relevancia de estas suposiciones a la hora de distinguir las competencias de la ciencia y de la filosofía a la hora de abordar los ámbitos explicativo-descriptivo y el justificativo de lo moral (Apel, 1985; Korsgaard, 2000; Cortina, 2007, 2012a; Evers, 2010). Tanto el establecimiento intersubjetivo de creencias, del que participa la ciencia en la comunidad científica, como el establecimiento intersubjetivo de lo normativo, en cuyos procedimientos y cuya argumentación la filosofía moral interviene de forma necesaria, se cumple el primero de los presupuestos, a saber, la suposición de un mundo objetivo más allá de los sujetos. Que también Apel reconoce como condición de posibilidad para la argumentación (Apel, 1985: 374). Respecto a este mundo, el establecimiento de creencias en el ámbito científico es deudor en la medida en que la verdad de estas depende de la relación que mantienen con ese mundo. Sin embargo, el reconocimiento de un mundo objetivo más allá de los sujetos no guarda esta relación con el establecimiento de lo normativo. Si bien existe el supuesto de un mundo compartido por los interlocutores, el establecimiento de lo normativo no depende del valor veritativo y la adecuación de los enunciados resultantes de los actos comunicativos de habla con ese mundo, sino del carácter público que adoptan las razones privadas al ser exteriorizadas en los contextos comunicativos de establecimiento de normas. Entran aquí en juego el resto de las suposiciones pragmático-formales. La suposición recíproca de racionalidad y responsabilidad de los interlocutores, la incondicionalidad de validez de los enunciados resultantes de los actos comunicativos de

habla y el discurso racional como foro último e irrefragable constituyen los elementos que permiten establecer lo normativo, en tanto que asentimiento reflexivo intersubjetivo, racional y establecedor de normas, que habrán de ser seguidas bajo supuestos de racionalidad y responsabilidad. Podría decirse, en pocas palabras, que lo relevante para el establecimiento dialógico y discursivo de la normatividad no corresponde a cómo *es* el mundo objetivo compartido, sino cómo *debe ser* bajo criterios de intersubjetividad, racionalidad y responsabilidad. Así los actos de habla comunicativos permiten ocuparse de la tarea justificativa de lo normativo, que orbita en torno al deber ser. Por su parte, aquellos actos comunicativos que se ocupan de establecer enunciados con pretensión de verdad acerca del mundo se ocupan del ámbito descriptivo y explicativo, que orbita en torno a lo que de hecho es. La verdad de una creencia respecto al mundo externo dependerá de si el enunciado representa lo que es el caso, mientras que la validez de una norma dependerá de las condiciones de intersubjetividad, racionalidad y responsabilidad en las que se asiente a la norma, o dicho de otro modo, del procedimiento. Si bien en la ciencia el procedimiento también resulta clave, en última instancia la verdad de los enunciados no depende de este, aunque sí lo haga la vigencia y la validez de las hipótesis y teorías.

Como hemos señalado, esto no significa que en el establecimiento de lo normativo no se vea implicada la suposición de un mundo objetivo común a los interlocutores. Los enunciados con pretensión de verdad sobre el mundo pueden resultar relevantes para informar a la ética (Quintelier et al., 2011; Kurtz y Koepsell, 2007; Rosenberg, 2000; Binmore, 2005; Pigliucci, 1987; Sober, 1994; Flanagan, 1996; Casebeer, 2003; Harman, 2012: 560) y tener en cuenta la situación concreta desde la que se produce la argumentación normativa (Apel, 1985; Habermas, 2006; Cortina, 2007, 2009, 2012a), de este modo la búsqueda de la verdad se halla imbricada con las intervenciones prácticas, para las cuales resulta necesario suponer la existencia de un mundo externo, en el cual la norma pretende ejercer su potencial regulativo. Dicho de otro modo, si bien la fuerza normativa de los enunciados regulativos de la praxis no depende de su correspondencia con el mundo en términos de verdad, han de atender a este en la medida en que esa praxis se produce en ese mundo compartido por los interlocutores. Asimismo, las razones que permitan el asentimiento de una norma, si bien no derivan un enunciado normativo a partir de enunciados informativos sobre el mundo, sí se sirven de estos para aportar razones a favor de aquel en la medida en que sean precisos ciertos conocimientos sobre

el mundo para formar argumentos acerca del deber moral. Es decir, la ciencia no fundamenta la moral, pero puede aportar conocimiento que ayude a la filosofía moral o ética en su tarea de fundamentación (Cortina, 2007, 2012a; Moore, 1983; Teehan y diCarlo 2004).

Los sujetos que forman parte de una comunidad de lenguaje determinada experimentan además una precomprensión gramatical del mundo que se adquiere por socialización, no tomando los elementos del mundo como objetos neutrales. Esta precomprensión supone una interpretación tal y como vimos con Apel (Lenk, 1982, 1988; Apel, 1985: 364; Conill, 2006; Gadamer, 1991). En lo que concierne a la normatividad también esta es heredada y configurada en términos de esta precomprensión. En ambos casos el lenguaje y la intersubjetividad que este permite en ámbitos comunicativos se constituyen como elementos abridores de mundo que permiten además establecer lo normativo, primero desde una precomprensión y después desde una crítica dialógico-discursiva en actos comunicativos de habla. Está implícita aquí una distinción de dos formas de lo moral a la que ya hemos atendido, la de moral como estructura y moral como contenido (Aranguren, 1994; Hauser, 2008; Cortina, 2012a). Los individuos humanos son estructuralmente morales en tanto que sujetos lingüísticos, racionales, autorreflexivos y comunicativos, estructuras que vienen marcadas a nivel neurobiológico y que suponen la base de lo moral (Darwin, 1979; Ayala y Cela Conde, 2001; Ayala, 2001; Damasio, 2008; 013; Waal, 1997, 2007; Brosnan y Waal, 2003; Tomasello y Vaish, 2012; Prinz, 2001; Nichols, 2005; Joyce; Hauser, 2008; Aranguren, 1994; Cortina, 2012a). Los contenidos morales son adquiridos durante el proceso de endoculturación siendo siempre históricos y contingentes, dependiendo además de la precomprensión gramatical que los diferentes lenguajes inculcan a los individuos, y serán siempre susceptibles de revisión y corrección, que será ejercida de forma pública en los contextos comunicativos con el fin de poder generalizar y aplicar intersubjetivamente las conclusiones resultantes de la crítica.

Los miembros de una determinada comunidad de lenguaje experimentan todo lo que les sale al encuentro en el mundo a la luz de una precomprensión «gramatical» adquirida por socialización, no como objetos naturales. La mediación lingüística de la referencia al mundo explica la retroconexión de la objetividad del mundo —objetividad que en la acción y en el habla se presupone— con la intersubjetividad del entendimiento entre participantes en la comunicación. Cualquier hecho que yo *enuncio* de un objeto, debe ser *afirmado* y, en su caso, *justificado* frente a otros que pueden contradecirme. La específica necesidad de interpretación surge de la

circunstancia de que ni si quiera en un uso descriptivo del lenguaje podemos prescindir de su carácter «abridor de mundo». (Habermas, 2006: 42-43)

Varias cuestiones se hacen aquí patentes. En primer lugar la pretensión de objetividad de la ciencia, en tanto que ejerce una labor descriptiva del mundo, se revela como sujeta a las interpretaciones tanto subjetivas como intersubjetivas que parten ya de esa precomprensión gramatical, la legitimidad de la ciencia no descansa en su objetividad neutra sino en la tendencia a esa objetividad ejercida en los actos comunicativos que tratan de establecer qué puede decirse acerca de un mundo objetivo compartido por los interlocutores. Como ya hemos podido ver, la existencia de este mundo es un presupuesto. En segundo lugar, se pone de manifiesto la posibilidad de escapar a la inconmensurabilidad en la medida en que los actos de comunicación permiten una fusión de horizontes. Encontramos aquí que el planteamiento de Habermas presenta una dimensión de carácter hermenéutico, superando algunos problemas interpretativos mediante el ejercicio de la argumentación en el ámbito comunicativo de una comunidad de hablantes. El propio autor expone esta cuestión.

Podemos superar reflexivamente nuestras diferentes situaciones hermenéuticas de partida y llegar a concepciones intersubjetivamente compartidas sobre la materia discutida. Es lo que Gadamer describe como «fusión de horizontes». (Habermas, 2006: 50).

Hasta aquí hemos abordado la suposición de un mundo objetivo compartido. Atenderemos ahora a la suposición recíproca de racionalidad de los interlocutores, en la que entran en juego la libertad y la responsabilidad. La relación entre razón y responsabilidad puede ilustrarse de varias maneras. Habermas afirma que el hecho de no poder dar cuenta ante otros de por qué se ha actuado de cierta manera hace sospechar que no existían razones para ello (como ya sugería Haidt al hablar del desconcierto moral: Haidt, 2001), lo cual supone haber obrado de forma irresponsable. Así, la responsabilidad se vincula al hecho de actuar según los dictados de la razón. Ahora bien, aún hace falta incluir otro elemento en la fórmula, la libertad. Solo si el sujeto actúa libremente puede achacársele la responsabilidad respecto a la acción (Ayala 2001; Cortina, 2012a: 149-201; Evers, 2010: 74-112). La razón deviene así práctica en la medida en que se orienta hacia la acción y trata de regularla, convirtiendo al actor en sujeto responsable en la medida en que actúe libre y racionalmente. La responsabilidad, dice Habermas, consiste en términos generales en la capacidad de un actor para orientar sus acciones por pretensiones de validez. Estas han de poder justificarse con razones ante interlocutores en

contextos de acción comunicativa. Esta concepción de las relaciones entre libertad, razón y responsabilidad plantea que los seres racionales son seres que actúan por criterios de racionalidad, haciéndolo libremente y convirtiéndose así en seres responsables de sus acciones. Las pretensiones de validez que guían la acción de los actores morales se suponen sujetas a razones que puedan justificar los motivos de acción ante interlocutores que exigen esta justificación.

Esta facultad de ser responsable no deja de ser, asegura Habermas, una suposición contrafáctica, pues en la vida cotidiana muchas expresiones y acciones vienen marcadas por motivos bien diferentes a las razones. Una ética de la razón cordial atiende a algunos de esos motivos, los correspondientes a factores emocionales implicados en la decisión y la acción moral, conocer y educar los motivos emocionales de la acción abre la posibilidad de ser responsables también en ese nivel, que puede complementar a la razón, pero no es exclusivamente racional (Cortina, 2009).

En tercer lugar, hemos de atender a la suposición de la incondicional validez de los enunciados resultantes de los actos comunicativos de habla. Habermas señala —en contraposición con la propuesta de Pierce (1965)— que el asentimiento por parte de sujetos racionales ante un enunciado con pretensión de verdad se debe a que el enunciado es verdadero, no es verdadero por el hecho de que se dé tal asentimiento mediante un consenso alcanzado en condiciones ideales. Un enunciado con pretensión de verdad idealmente justificado puede acabar revelándose a la postre como falso, esto pone de relieve que el asentimiento intersubjetivo en actos comunicativos permite tomar un enunciado como verdadero, en tanto que válido y vigente, pero no garantiza la correspondencia de lo que el enunciado predica con lo que es el caso en el mundo objetivo compartido. Los procesos de argumentación máximamente inclusivos que Habermas plantea no posibilitan la comprobación y el cercioramiento definitivo e indudable de la verdad de los enunciados ante los que se asiente, pero al mismo tiempo constituyen la mejor baza de la que disponemos para proponer enunciados con pretensión de verdad referidos al mundo externo. La solución no pasa por desestimar los procesos dialógicos de argumentación, sino por reconocer que la intersubjetividad del asentimiento resultante se produce en términos de “aceptabilidad racional”, y no en términos de “verdad”. Llegamos con ello a la cuarta suposición que indica Habermas, la que atiende a la razón como capacidad de principios puesta en relación con el discurso racional y argumentativo

que supone el foro último e irrebable de toda posible jurisdicción, a pesar de su carácter revisable y actualizable desde la superación de las actuales formas de entendimiento.

Estas consideraciones respecto al tercer y cuarto supuesto, y también algunas de las que aportamos respecto al primero, conciernen especialmente a los enunciados asertóricos del conocimiento sobre el mundo, con pretensión de validez, que aspira a la verdad del enunciado como adecuación a lo que es el caso en el mundo objetivo (Habermas, 2006: 62). La ciencia empírica propone enunciados de este tipo, presentando constantes revisiones y propuestas de teorías, que se producen en el marco comunicativo y argumentativo de la comunidad científica. Por otra parte, Habermas muestra interés en otro tipo de enunciados, respecto a los cuales merece la pena investigar cómo las premisas del concepto kantiano de razón práctica permiten pretensiones de validez incondicional, hablamos aquí de los enunciados morales correctos, y con ciertas reservas, dice Habermas, también los jurídicos.

El caso de los enunciados con pretensión de verdad acerca del mundo objetivo presenta diferencias cruciales en relación al caso de los enunciados con pretensión de validez normativa que permiten regular la conducta moralmente. Aquellos aspiran a una referencia acertada respecto a un mundo objetivo en términos de verdad. En estos, sin embargo, mediante actos regulativos de habla, la referencia se produce no respecto a un mundo objetivo, sino respecto a acciones en relación a las cuales los actores se sienten obligados. El contexto social y el grupo del que forman parte los individuos inculcan en este una serie de valores, prácticas y normas comunes que son reconocidos intersubjetivamente en el seno del grupo o apoyados implícitamente por el hábito y la costumbre (Ayala y Cela Conde, 2001; Hauser, 2008; Cortina, 2012a, 2017; Evers, 2010). De hecho, y a diferencia de lo que concierne a los enunciados descriptivos sobre el mundo, la validez en los casos en que hay una pretensión de corrección varía según el trasfondo cultural y social desde el que se legitima tal pretensión. No obstante, asegura Habermas, los imperativos morales pretenden validez absoluta, es decir, pretenden reconocimiento universal. Este se posibilita mediante actos comunicativos de habla que permitan poner en común todas las razones y asentir a las que presenten los mejores argumentos. El discurso dialógico permitirá revisar los criterios de vigencia de los elementos normativos, facilitando el asentimiento y el rechazo de estos, según proceda, mediante la reflexión intersubjetiva y el intercambio argumentativo de razones. Los

participantes en estos actos de comunicación se mueven sobre el presupuesto de que el resto de participantes son agentes libres, racionales y responsables que pueden vincular su voluntad a máximas, asumiendo obligaciones y deberes. Todo ello bajo el presupuesto contrafáctico de la ampliación de los contextos sociales que posibilitará una fusión de horizontes gracias a la cual el asentimiento ante las normas racionalmente argumentadas podrá producirse de forma universal.

Los participantes en los actos comunicativos no consideran, o al menos no deben considerar, que lo normativo forma parte del mundo objetivo compartido por todos los interlocutores. En su lugar lo normativo, que no se halla en el mundo objetivo, se produce como resultado del intercambio de razones que permite la reflexión y el asentimiento intersubjetivo ante criterios de normatividad que responden a expectativas de conducta generalizadas que pueden apoyarse argumentativa y públicamente. Así lo normativo se establece como resultado del proceso de autorreflexión y asentimiento intersubjetivo que supone el diálogo discursivo, y solo entonces pasa a existir. Los enunciados morales, y todo enunciado normativo en general, no refieren a un mundo objetivo que existe con independencia del enunciado, sino que permiten, como ya apuntaba Korsgaard, la creación de entidades intrínsecamente normativas como resultado de procedimientos adecuados de establecimiento de la normatividad (Korsgaard, 2000). Procedimientos en los que el agente moral o una comunidad de agentes morales se constituyen como la fuente misma de la normatividad, estableciendo entidades normativas, las normas, mediante asentimiento reflexivo y ejercicio de la autonomía, que se ejerce de forma colectiva e intersubjetiva en el caso de las éticas dialógicas. Frente al punto de vista de la tercera persona, que se produce en la perspectiva del observador propia de las ciencias, y completando el insuficiente punto de vista de la primera persona a la hora de establecer lo normativo, entra en juego el papel de la segunda persona, en relación a la cual la reflexión se produce en actos comunicativos y argumentativos, también en relación a esa segunda persona se tienen expectativas, así como obligaciones, en lo que concierne al cumplimiento de normas. Teniendo en cuenta este carácter intersubjetivo y comunicativo del procedimiento, que permite establecer normas, y la fuerza normativa de estas, carece de sentido tan si quiera plantear que alguien pueda seguir una regla solipsistamente, para sí mismo. Esta cuestión es tratada pormenorizadamente por Apel, como hemos tenido ocasión de ver. La dimensión social de una regla hace que todo solipsismo tanto

metodológico como en el seguimiento de la regla carezca de sentido (Habermas, 2006: 75; Apel, 1985: 374).

En la medida en que la universalidad de la norma depende de una ampliación de la comunidad de hablantes que permita una fusión de horizontes total, no podremos hablar de un asentimiento universal, sino que hemos de proponer procedimientos que marquen una tendencia a ese tipo de asentimiento. Habermas dice que podemos imaginar “los procesos de aprendizaje moral como la ampliación inteligente y el ensamblaje recíproco de mundos sociales que [...] todavía no se superponen suficientemente” (Habermas, 2006: 55). Esta idea guarda estrecha relación, aunque no se identifica, con la contradicción dialéctica que Apel señalaba entre la comunidad real de comunicación y la comunidad ideal de comunicación (Apel, 1985: 346-348). Este ensamblaje refiere a la fusión de horizontes que los integrantes de diferentes contextos sociales van experimentando a medida que se incorporan como interlocutores en actos de comunicación que pretenden establecer la corrección de enunciados normativos. Las perspectivas iniciales, que son por defecto etnocentristas (Evers, 2010), y que extienden sus raíces en el suelo social en el que han germinado, van descentralizándose y proyectándose a medida que la argumentación discursiva y dialógica pone en la esfera pública los motivos racionalmente argumentados que proponen como guías de la acción miembros de otras sociedades o grupos. Así el intercambio comunicativo de razones conduce a un entrecruzamiento de perspectivas diversas y de orientaciones valorativas diferentes, que convergen en la medida en que los interlocutores se ven forzados a asentir ante la fuerza de los mejores argumentos, bajo el supuesto de racionalidad, libertad y responsabilidad atribuidos a todos los participantes en los actos de habla comunicativos, que tienen como objetivo alcanzar acuerdos en torno al establecimiento de lo normativo.

Habermas es consciente de que los estándares bajo los cuales algo puede ser considerado como un buen o un mal argumento pueden variar mucho. Por ello trata de garantizar bajo la propuesta de una situación ideal de habla unas condiciones exigentes y determinadas bajo las cuales la aceptabilidad racional puede producirse de modo que unas razones puedan ser aceptadas contra toda objeción posible. La situación ideal de habla ha de presentar unas presuposiciones pragmáticas que son inherentes a toda práctica argumentativa —idea compartida por Apel (Apel, 1985) y Cortina (2007)—. Según Habermas, estas son las siguientes: la argumentación ha de tener un carácter público e

inclusivo, de modo que el asentimiento se produzca como resultado de deliberación pública en la cual es susceptible de participación todo interlocutor real o virtual; se ha de dar una igualdad de oportunidades en el ejercicio de las facultades comunicativas, de tal modo que todos puedan expresarse en torno a la materia tratada; se ha de evitar todo engaño e ilusión, para ello los participantes han de hacer ejercicio de sinceridad y defender solo posturas que suscriban; por último, la argumentación se ha de producir sin la presencia de coacción. Estos presupuestos implican un universalismo en relación a las cuestiones práctico-morales, exigiendo a su vez por parte de los participantes una actitud crítica con los demás y consigo mismo y una apertura hermenéutica que predisponga a los individuos hacia la comprensión que los otros tienen de sí mismos, del mundo y de los criterios morales que deberían guiar la acción.

Existen objeciones que pueden ponerse ante una propuesta de este tipo, ya que muchas veces los participantes en la argumentación actúan por motivos que no son en absoluto racionales, o que responden a una racionalidad estratégica. La situación ideal de habla elude estas críticas en la medida en que es una idea regulativa, y no una propuesta de ejercicio real de establecimiento de lo normativo. Recordemos que Apel advierte de cómo en la comunidad real de comunicación se producen interferencias por el hecho de que los individuos argumentan desde su posición concreta, lo cual les aleja de la situación ideal de habla, que ha de servir como ideal regulativo (Apel, 1985). Precisamente el carácter racional y cooperativo que la situación ideal de habla propone de manera contrafáctica sirve para regular esas infracciones, tanto las no racionales como las estratégicas, de las expectativas de conducta que pueden ponerse sobre el tapete mediante argumentación pública, orientadas a facilitar una convergencia de perspectivas que permita establecer tales expectativas como pautas de conducta regulativas con fuerza normativa, en tanto que se asienta ante ellas. Además, la situación ideal de habla en tanto que idea regulativa marca una tendencia a la extensión e inclusión de participantes en el diálogo, que tendría como último término una deliberación y un establecimiento de lo normativo en términos universales, de forma similar a cómo se resolvería la contradicción dialéctica entre la comunidad de argumentación ideal y la comunidad de argumentación real según Apel (Apel, 1985: 346-348). Una ética de la razón cordial, además, permite tomar en consideración los aspectos emocionales y estimativos implicados en el diálogo y la decisión moral, evitando así que supongan una interferencia, el hecho de considerar

el diálogo como un acto emotivo-racional elude los problemas derivados de una perspectiva de corte meramente racionalista (Cortina, 2009).

Esta dinámica —incorporada en la práctica argumentativa— de un descentramiento continuo y progresivo de las propias perspectivas interpretativas impulsa especialmente los discursos prácticos en los que no se trata ya de cerciorarse sobre las pretensiones de verdad, sino de la construcción razonada y la aplicación de normas morales (y jurídicas).

La validez de tales normas consiste en el reconocimiento universal que las normas merecen. (Habermas, 2006: 59)

Se recoge aquí un aspecto fundamental del establecimiento de lo normativo, sobre el que volveré más adelante. Las normas morales no se establecen en función de elementos intrínsecamente normativos del mundo, como ya denunciaba Korsgaard (2000: 53), es decir, no pueden fundamentarse desde el paradigma naturalista (Buchanan, 2009; Apel, 1985; Habermas; Cortina; García-Marzá, 2016; Aranguren, 1994; Teehan y diCarlo 2004; Gould, 1999; Hume 1988, 1953; Moore 1983; Salles 2013, 2016, 2018; Evers 2010; Conill, 2012), sino que se construyen mediante razonamiento reflexivo y se validan mediante asentimiento, siendo el consenso el tipo de asentimiento que se produce intersubjetivamente en comunidades de comunicación. La norma moral constituye el resultado construido de procedimientos de establecimiento de lo normativo, pasando a ser esta intrínsecamente normativa, pero también revisable y susceptible de corrección. Solo en este sentido cabe hablar de cierto tipo de realismo, en la medida en que el agente moral construye entidades normativas, las normas, que son reales en virtud del consenso que establece su fuerza normativa (Korsgaard, 2000).

Respecto a las pretensiones de validez que aspiran a convertirse en norma bajo asentimiento en contextos de comunicación, es decir mediante consenso, cabe tomarlas en consideración, por un lado, en tanto que pretensiones y, por otro, en tanto que aspiran a ser reconocidas de forma incondicionada. Como pretensiones quedan sujetas al reconocimiento intersubjetivo que supone un consenso que asiente. En tanto que se pretende una validez incondicionada estas pretensiones apuntan más lejos que al mero consenso fáctico, siempre revisable y susceptible de corrección, su aspiración en último término, bajo una presuposición idealizante, consiste en que a largo plazo se pondrán sobre la mesa todas las informaciones y razones relevantes, lo cual haría del asentimiento ante una norma algo incondicionado. Así, se puede afrontar una intuición trascendental

de que la objetividad puede encontrar un fundamento último e irrebasable en la intersubjetividad que permiten los actos lingüísticos comunicativos (Habermas, 2006: 82). Aquí resulta especialmente relevante la idea, señalada por Apel, de que la realización de la comunidad ideal de habla es presupuesta en la comunidad real de habla.

Naturalismo blando, perspectivas epistémicas y perspectivas metodológicas.

En la tercera parte de *Entre naturalismo y religión* Habermas plantea dos ideas que resultan clave para nuestra investigación en torno al método en neuroética. Atenderemos especialmente al primero de los ensayos que componen esta tercera parte, en el que plantea de forma explícita, aunque un tanto confusa, un concepto de naturalismo que acuña como *naturalismo blando*. Profundizaremos en este concepto y trataremos de aclararlo. En relación con este concepto de naturalismo Habermas plantea una dualidad de perspectivas metodológicas y epistémicas, a las que también atenderemos.

Con el objetivo de abordar la cuestión del libre albedrío y el determinismo Habermas ha de ocuparse necesariamente de la relación entre el naturalismo y el espíritu, especialmente cuando el intento de naturalización adopta actitudes totalizadoras. No profundizaremos ahora en la cuestión del libre albedrío, a la que ya tuvimos ocasión de atender (Evers, 2010; Cortina, 2012a, 2012b; Libet, 1999; Libet et al., 1983; Farah 2012; Illes y Sahakian 2012; Kant, 1992, 2000; Roskies, 2010), sino en cómo esa naturalización se configura en torno a elementos propios del espíritu, muchas veces inabarcables para el naturalismo. En relación a estos elementos iremos esbozando qué cuestiones conciernen al método en neuroética.

Habermas denuncia las estrategias reduccionistas llevadas a cabo desde algunas esferas de las ciencias naturales. En lo que concierne a la libertad o libre albedrío y al *yo*, la neurobiología ha tratado de explicar ciertos procesos y percepciones mentales partiendo únicamente de condicionamientos fisiológicos observables. Se reduce así el sujeto de acción del *yo* al cerebro, y la libertad puede plantearse como una mera ilusión perceptiva ante la posibilidad de un determinismo configurado en la conexión causal de procesos neuronales sujetos a las leyes de la naturaleza. Habermas se pregunta si esta concepción determinista está fundada en las ciencias de la naturaleza, o si más bien es un elemento que forma parte de una imagen del mundo de carácter naturalista, resultado de una interpretación de ciertos conocimientos aportados por las ciencias naturales.

Habermas se inclina por esta segunda opción. Declara aquí que desea continuar el debate sobre libertad y determinismo como “una confrontación acerca del *modo correcto* de naturalización del espíritu” (Habermas, 2006: 160). La cuestión del método, entendido como modo correcto o adecuado de investigación, se halla presente en esta afirmación.

Habermas expone las partes del ensayo. Planteará en primer lugar que los programas de investigación reduccionista solo pueden resolver las dificultades propias de un dualismo de perspectivas y juegos de lenguaje incurriendo en posturas propias del epifenomenalismo, lo cual sería deseable evitar. Según el epifenomenalismo la relación entre materia y mente se da en un solo sentido, unidireccionalmente, siendo la materia capaz de actuar sobre la mente, pero no inversamente. En esta cuestión Cortina se muestra especialmente crítica. Expone el caso de las razones y su influencia en la acción, estas no funcionan como condiciones causalmente eficientes para la acción, pero ciertamente influyen en ella. Las razones no son estados físicos observables, sino que se ubicarían en la esfera de lo mental. Si las razones pueden identificarse con la mente y no con un estado físico, e influyen en la acción, entonces el epifenomenalismo no será sostenible.

Quien mantiene esta posición debería explicar por qué las razones influyen en las acciones, por qué cuando un sujeto se propone llevar a cabo una acción el tipo de razones que tiene para realizarla tiene influencia en ella. (Cortina, 2012a: 193)

Cortina aporta algunos datos que respaldan su tesis. Entre ellos presenta los experimentos de Peter Gollwitzer (1999), que muestran la relación entre las intenciones y la realización de las acciones que dichas intenciones prometen. El alto grado de coincidencia entre la intención y la realización de la acción son un indicador, dice Cortina, de la influencia de los estados mentales en la acción. También atiende a Searle (1992, 1996, 2006), quien acertadamente denuncia que el epifenomenalismo va en contra de todo lo que sabemos sobre la evolución. El desarrollo evolutivo de la mente es un proceso muy complejo, cuyo resultado, la mente, supone un alto coste para los individuos que la poseen. Carecería de sentido que los estados mentales no pudiesen influir en la materia y en la acción, quedando reducidos a epifenómenos. La resolución de problemas, la predicción del futuro, el manejo de razones, hacer balance de estas, etc., son capacidades mentales que cumplen una función clave en la conducta, solo tiene sentido que se hayan desarrollado evolutivamente si la relación materia-mente es bidireccional. Por último, Cortina critica la pretensión de que la conciencia del *yo* sea una ilusión y la conciencia de

la libertad un epifenómeno, tal como afirma Wolf Singer (2004). Si bien no es posible identificar un correlato neurobiológico del *yo*, eso no significa que no exista ese *yo* o que sea una ilusión. El *yo* puede ser una construcción, pero eso no implica que sea una ilusión, como indica Habermas.

En segundo lugar, volviendo a Habermas, este analizará las raíces antropológicas de ese dualismo, el cual no significa una exclusión de una visión monista de la evolución natural. Respecto a ello argumentará que la relación cerebro-espíritu se produce como resultado de una reflexión filosófica y no como resultado del conocimiento de las ciencias naturales.

Antes de desarrollar los puntos que señala, Habermas declara su intención de argumentar a favor de lo que denomina un *naturalismo blando*. Introduce en este punto una descripción de este.

Yo defiendo un naturalismo «blando», no cientificista. Según esta concepción, solo es «real» lo que puede ser expresado en enunciados verdaderos. Pero la realidad no se agota en el conjunto de enunciados regionalmente limitados que, según los estándares actuales, cuentan como verdaderos enunciados científicos de la experiencia. (Habermas, 2006: 161).

Analicemos la cita con detenimiento. En primer lugar, Habermas considera que el naturalismo blando no es cientificista, es decir, no otorga a la ciencia una potestad omniabarcante, no cree que todo sea abarcable desde el prisma científico. En segundo lugar, afirma que lo real es aquello que puede ser expresado en enunciados verdaderos, es decir, un enunciado verdadero es aquel que expresa la realidad en tanto que refiere a ella, además enuncia aquello que es el caso en ese mundo objetivo compartido del que hemos hablado anteriormente. A ello añade que la realidad no se agota en el conjunto de enunciados que están limitados por presentar un carácter regional, en tanto que se hallan sujetos a estándares actuales, susceptibles de ser modificados o sustituidos, a pesar de que esos enunciados cuenten actualmente como enunciados científicos de la experiencia que refieren al mundo objetivo.

Es decir, el naturalismo blando al que Habermas refiere pone de relieve que los enunciados que la ciencia toma a día de hoy como verdaderos son susceptibles de revisión, corrección y sustitución, encontrándose sujetos a unos estándares relativos a su aceptación que pueden variar histórica y socialmente. Por ello considera que lo real no se

agota en estos conjuntos de enunciados, no puede aspirarse a la articulación de enunciados verdaderos definitivos más que como una idea regulativa del ideal científico, pues todo enunciado que se defiende como pretendidamente verdadero queda fácticamente sujeto a revisión y es susceptible de ser a la postre falso, aún si tiene vigencia actualmente como verdadero. La comunidad científica compone el contexto argumentativo y dialógico que permite el establecimiento de enunciados verdaderos, entendiendo esta verdad en sentido dialógico y como resultando del consenso en la comunidad científica, entendida como comunidad de comunicación. La verdad de tales enunciados, por tanto, será vigente hasta que nuevos argumentos de los participantes pongan fin a esa vigencia. En este sentido Habermas sostiene que la pretensión de un *naturalismo duro* que atienda describa, refiera y acceda directamente a los hechos no puede sostenerse, pues no existe más acceso al mundo objetivo que el que permite la intersubjetividad resultante de la argumentación discursiva y dialógica. Ni tan siquiera la experiencia permite este acceso, ya que las teorías e hipótesis científicas serán siempre interpretaciones de los hechos empíricos, y no un reflejo directo de estos (Habermas, 2006; Conill, 2006; Gadamer, 1991). Lo que Habermas parece apuntar al hablar de naturalismo blando es el carácter revisable y falible de los enunciados científicos, que son enunciados que tienen pretensión de verdad respecto al mundo objetivo compartido. La crítica en relación a este concepto es posible (Ortega Esquembre, 2018), en aras de mantener la coherencia con el resto de su pensamiento, Habermas debería haber usado términos más adecuados que el de *naturalismo* para referir a esta postura, aun si refiere a él como un naturalismo *blando* o *suave* (*soft*).

Frente a este naturalismo blando existen otras perspectivas metodológicas que no son ni incompatibles con él ni reductibles a él. Tal y como vimos en el apartado anterior, los enunciados con pretensión de validez no son el tipo de enunciados que pretenden referir y describir el mundo objetivo, sino que se destinan a orientar la acción. El riesgo del reduccionismo consiste en pretender adjudicar a la primera de las perspectivas metodológicas las competencias de la segunda. En relación a ello, Habermas señala que es conveniente mantener la escisión epistémica entre perspectivas, diferenciando la perspectiva de la ciencia natural, y formas de saber complementarias, de la perspectiva de la filosofía moral. Esta distinción metodológicamente necesaria se ve socavada por la acción del reduccionismo (Buchanan, 2009; Korsgaard, 2000; Apel, 1985; Habermas; Cortina; García-Marzá, 2016; Aranguren, 1994; Teehan y diCarlo 2004; Gould, 1999;

Hume 1988, 1953; Moore 1983; Salles 2013, 2016, 2018; Evers 2010). A todo ello hay que sumar un problema añadido, que atañe a la cuestión de qué puede ser medido desde los experimentos propios de las ciencias naturales, subyace aquí la cuestión filosófica relativa a qué cosas pueden ser sometidas a medición (Habermas, 2006: 162). Habermas atiende aquí a los experimentos neurocientíficos de Libet en torno al libre albedrío, indicando los problemas de método que este tipo de experimentos conllevan.³³

Habermas pone de relieve el papel de las razones como motivos de la acción. Sin entrar en el debate libertad-determinismo ni en el problema de si está determinado el propio proceso de dar y pedir razones, así como la influencia motivadora de tales razones sobre la acción, podemos ver que la perspectiva de las ciencias naturales es insuficiente para abordar dichos procesos. La *perspectiva del observador*, propia de las ciencias naturales, no puede prescindir de la *perspectiva del partícipe*, que es constitutivamente subjetiva, Evers afirma en este sentido que la conciencia es ontológicamente subjetiva (Evers, 2010). El observador ha de recurrir a un lenguaje mentalista cuando trata de describir el proceso discursivo, utilizando predicados como opinar, convencer, afirmar o negar. El uso de un lenguaje meramente empirista tendría que eliminar toda referencia a las posiciones proposicionales adoptadas por los sujetos, esto supondría tomar el proceso discursivo como un acontecimiento natural que se produce a “espaldas de los sujetos”, ilustra Habermas (2006). En este punto la crítica se orienta hacia las posturas de Peter Bieri (2002) y Davidson (2001, 2004), quienes plantean un monismo ontológico según el cual los motivos racionales y los elementos mentales son la misma cosa que las causas, existiendo un correlato entre mente y cerebro. Si bien Habermas reconoce que todo proceso mental tiene base biológica y cerebral y depende de esa base, considera que la identificación entre motivos y causas es una reducción inapropiada. Podemos extraer la siguiente conclusión, que nos resulta valiosa para el tema que aquí abordamos: la perspectiva del observador no puede prescindir de la perspectiva del partícipe, a menos que incurra en el epifenomenalismo, si así lo hiciera quedarían fuera de nuestro campo de visión ciertos aspectos relevantes del espíritu.

En este punto, Habermas nos advierte de los errores que Kant y Descartes cometieron, y que debemos evitar. Kant manejó un concepto de libertad inflacionario

³³ Hemos llevado a cabo un análisis pormenorizado de la cuestión en el segundo epígrafe del Capítulo 4.

inteligible y sin origen, separado de todas las conexiones empíricas, un concepto de libertad absoluto. Este concepto de libertad es superado por las propuestas dialógicas, que comprenden la libertad situada y puesta en ejercicio en contextos comunicativos concretos (Apel, 1985; Habermas, 2010), y que además se halla vinculada a aspectos emocionales y no solo racionales, tal y como indican las propuestas que suscriben la idea una razón encarnada o cordial (Cortina, 2009; Conill, 2015b, 2016; González Esteban, 2013). Descartes, por su parte, planteó una ontología dualista, asociada a un dualismo de perspectivas. Dar cuenta de esa doble ontología resulta problemático. Habermas señala que los agentes no están fuera del mundo y tanto su libertad —en caso de que exista el libre albedrío— como las dos perspectivas que pueden adoptar, del partícipe y del observador, están enraizadas en el organismo, en la naturaleza y en su historia vital. Un dualismo ontológico ha de ser desestimado en favor de un monismo ontológico. Por ello la libertad no puede plantearse en sentido absoluto, sino condicionado (Cortina, 2012a: 187-192). En cuanto al dualismo de perspectivas, este se ancla de forma última en un mismo sustrato, el natural, de ahí que trasladar esa dualidad de perspectivas a una ontología dual sea un error.

El dualismo de perspectivas explicativas y juegos de lenguaje es difícil de superar, y compone el corazón de los principales problemas de método en la interacción de las ciencias naturales con otros ámbitos como la filosofía. Por ello esta cuestión resulta especialmente interesante a la hora de elaborar una neuroética crítica de la razón cordial, que será constitutivamente interdisciplinar.

Tal dualismo epistémico tiene, ciertamente, un sentido metodológico, no ontológico. Sin embargo, hasta ahora sigue estando poco claro cómo éste puede ponerse en concordancia con una concepción monista del universo que se ajuste a nuestra necesidad de tener una imagen coherente del mundo. (Habermas, 2006: 170)

Como puede verse, el problema reside en la imposibilidad de hacer converger en una sola dos perspectivas diferentes que tratan de dar cuenta de una misma realidad, en virtud de ello encontramos un monismo ontológico con el que nos podemos relacionar metodológicamente de forma dual. Esto nos remite al distanciamiento del sujeto respecto al objeto propio de las ciencias naturales, del que nos hablaba Apel, que no podía producirse en las ciencias humanas, en la medida en que no puedan participar plenamente de la perspectiva del observador (Apel, 1985). La solución adecuada a este problema no

es la búsqueda de una convergencia de ambas perspectivas en una sola, sino el reconocimiento de que una compatibilidad de esas dos perspectivas es posible, estableciendo qué competencias corresponden a cada una de ellas (Korsgaard, 2000; Apel, 1985; Habermas, 2006; Cortina, 2012a; Evers, 2010; Gould, 1999; García-Marzá, 2016).

Quienes sostienen posturas de naturalismo reduccionista o de cientificismo argumentan que ambas perspectivas no comparten una igualdad de derechos (Churchland, 2013; Greene; Gazzaniga, 2006; Mora, 2007; Casebeer, 2003). Del mismo modo que un fenómeno subjetivo como el calor puede ser explicado exitosamente en términos objetivos desde la física, “la interacción entre espíritu y cerebro podrían tener la respuesta correcta en planteamientos de investigación que solo se fíen de las duras explicaciones causales, mandando las blandas explicaciones racionales al mundo ilusorio de la psicología de la cotidianidad” (Habermas, 2006: 170). Responde Habermas que de ello se derivaría considerar toda praxis justificadora como epifenoménica, lo cual es resultado de un planteamiento reduccionista de la investigación. Si el espacio de las razones no desempeña desde el punto de vista neurobiológico ningún papel causal, como motivos de la acción o la decisión, sería un derroche por parte de la naturaleza el desarrollo filogenético y ontogenético del ámbito de las razones. Lo que sostiene Habermas es que los procesos de juicio vienen siempre ligados a razones. Las propias funciones biológicas del proceso de dar y pedir razones exigen considerar las razones como algo más que epifenómenos. Algo que resulta evidente ante el hecho de que las razones influyen la acción (Cortina, 2012a).

Habermas concluye que la apuesta más acertada es considerar las dos perspectivas, la del observador y la del partícipe, como herramientas que sirven para abordar de forma diferente cuestiones que pertenecen a un mismo mundo. Es decir, una dualidad de perspectivas no es incompatible con una visión monista del mundo. Respecto a ello Habermas reconoce a Rorty (2004) el mérito de algunas de sus observaciones:

En este sentido, Richard Rorty declara un desdoblamiento gramatical en nuestros vocabularios explicativos entre aquellos que dirigen su mirada hacia las causas observables y aquellos otros que la dirigen hacia los motivos racionales, resultado, todo ello, de una adaptación funcional de nuestra especie a dos entornos distintos, el natural y el social. La irreductibilidad de un juego de lenguaje al otro no debería intranquilizarnos más que el hecho de que un instrumento no sea sustituible por otro. (Habermas, 2003: 173)

Que se haga referencia al “desdoblamiento gramatical en nuestros vocabularios” es indicador de una cuestión que ya hemos apuntado. Tanto el ámbito explicativo de las ciencias que “dirigen su mirada hacia las causas observables” como el de las prácticas de establecimiento de lo normativo que “la dirigen hacia los motivos racionales” se ejercen desde vocabularios y actos comunicativos diferentes, con sus propias competencias, que permiten establecer enunciados con pretensión de verdad y enunciados con pretensión de validez, respectivamente. Cada tipo de enunciado responde a las exigencias de diferentes entornos, en el caso del entorno natural se busca un conocimiento de la verdad, en el caso del entorno social se buscan criterios de normatividad.

Tras exponer que ambos juegos de lenguaje o perspectivas son compatibles, aunque sean irreductibles el uno respecto al otro, es decir, que no puedan converger en una sola perspectiva o juego de lenguaje que los unifique, Habermas añade que, desde una perspectiva epistemológica, no está claro que pueda dejarse de lado ninguno de ellos. El reduccionismo naturalista pretende hacer tal cosa.

En primer lugar, Habermas señala que este dualismo epistemológico es irreversible, no podemos ya desprendernos de él, lo cual es un indicador de que los juegos de lenguaje y los modelos explicativos de cada una de las perspectivas no pueden reducirse los unos a los otros. No podría haber una traducción de un vocabulario mentalista a uno empirista sin una pérdida de contenido semántico. Del mismo modo no puede haber una traducción de un vocabulario normativo a uno empirista, sin una pérdida del mismo tipo. No podemos reducir todos los objetos mentales a meras cuantificaciones de procesos neurobiológicos. Conceptos morales como el bien, la justicia o la virtud, así como la obligación y el deber, y por supuesto las normas, difícilmente pueden someterse a medición, lo máximo a lo que se puede aspirar es a cuantificar los procesos cerebrales de los sujetos cuando abordan cuestiones relativas a tales elementos, pero tales procesos no son en absoluto aquello que es procesado por los mismos. No es posible hallar una correspondencia del carácter normativo de los conceptos morales con los procesos neurobiológicos subyacentes a la adquisición y uso de tales conceptos, pues lo normativo no refiere al mundo objetivo, sino que es resultado de poner en ejercicio ciertas facultades y procesos mentales en un contexto comunicativo mediante la interacción intersubjetiva bajo criterios de racionalidad.

Dice Habermas, en relación a la libertad, que todo intento de naturalizar el espíritu fracasa a la hora de llevar a cabo una traducción desde la perspectiva mentalista a la naturalista. También está abocado al fracaso todo intento de naturalizar lo normativo mediante una traducción naturalista. Respecto a tales objetos la ciencia natural no puede más que adoptar una postura agnóstica (Cortina, 2012a). Habermas arroja luz sobre la causa principal de esa intraducibilidad, asegura que “en la gramática de esos dos tipos de juegos de lenguaje se han instalado ontologías irreconciliables” (Habermas, 2006: 175). El error del reduccionismo cientificista es considerar que la ciencia, basada en una ontología naturalista y empirista, es capaz de abordar el estudio de objetos que son inaccesibles para una ontología de ese tipo, inaccesibles a la experiencia, puesto que rebasan la posibilidad de todo conocimiento empírico. La pretensión reduccionista de las ciencias naturales comete el grave error de querer estudiar algo para lo cual su perspectiva metodológica está constitutivamente impedida. Habermas se plantea si esta irreductibilidad entre las dos perspectivas no debería conducirnos a considerar al mundo desde ambas simultáneamente para así aprender algo de él adecuadamente. Refiere aquí a dos juegos de lenguaje o perspectivas distintas, una concebida para lo espiritual y otra para lo físico. De forma análoga, y atendiendo a lo expuesto en el apartado anterior, podemos reconocer la existencia de dos discursos diferenciados, el científico, que propone enunciados con pretensión de verdad, y el normativo, que propone enunciados con pretensión de validez normativa. Ambos se producen dialógicamente en una comunidad de hablantes que asienten ante los enunciados mediante el consenso, siendo todo enunciado revisable y potencialmente sustituible a la luz de una argumentación futura que ofrezca motivos suficientes para su derogación.

En las prácticas comunicativas y sociales encontramos un ejemplo de cómo la perspectiva del partícipe se hace necesaria, siendo imposible llevar a cabo tales prácticas meramente desde la perspectiva del observador. Incluso el quehacer científico el investigador, que adopta primeramente una postura objetivante desde la perspectiva del observador, ha de someterse luego al ámbito social y comunicativo de la comunidad científica, para defender sus tesis, adoptando una perspectiva de partícipe. El saber positivo fuerte de las ciencias naturales parece sobresalir respecto al saber débil de prácticas y entramados de sentido simbólicamente constituidos. Mientras que las ciencias naturales pueden apelar a la experiencia y la observación, presentando grandes logros explicativos y predictivos, así como una rica posibilidad de desarrollo técnico, que

aproveche el conocimiento científico, la otra forma de saber recurre a la interpretación y explicación de significados y de referencias prácticas, bajo control de preguntas y respuestas hipotéticas. Esto parece poner de relieve una superioridad de la perspectiva naturalista, lo cual ha conducido en algunos casos a la pretensión de sustituir ese saber hermenéutico débil por un saber positivo fuerte.

Ahora bien, no está tan claro que esta supuesta debilidad del discurso teórico y esta supuesta fortaleza del discurso normativo se den de este modo. La investigación científica opera argumentativamente en el ámbito de la comunidad científica, donde toda pretensión de conocimiento que se aporte desde la perspectiva del observador ha de ser presentada y defendido comunicativamente, lo cual supone utilizar recursos hermenéuticos, a lo cual hay que sumar que toda hipótesis y teoría es ya una interpretación de los hechos, y no un acceso directo a estos (Lenk, 1982, 1988; Apel, 1985: 364; Conill, 2006; Gadamer, 1991). Esto significa que en el ámbito científico la experiencia y la argumentación se revelan como dos partes dependientes la una de otra, necesarias para ver satisfechas nuestras expectativas de obtener algún tipo de conocimiento sobre el mundo externo compartido. Por tanto, el uso simultáneo de las dos perspectivas metodológicas se hace necesario (Habermas, 2006: 174-177, 213). En la medida en que las ciencias naturales han de participar de la perspectiva del partícipe no pueden ejercer la objetivación del mundo que pretenden. Ni pueden tampoco defender una fortaleza o relevancia mayor de su paradigma respecto a otros enfoques eminentemente interpretativos, sociales y prácticos, ya que la ciencia también presenta estos elementos. El resultado es pues, no un naturalismo duro que se ancla en un mundo objetivo, sino un naturalismo blando que refiere a ese mundo mediante enunciados aceptados como verdaderos provisionalmente en una comunidad de argumentación, la comunidad científica. Tanto en el discurso científico como en el normativo, así como en la vida cotidiana, el ser humano es observador y partícipe al mismo. Por ello, aun irreductibles la una la otra, las dos perspectivas metodológicas han de funcionar simultáneamente para una comprensión adecuada del mundo (Habermas, 2006; Evers, 2010: 71).

En relación a estas dos perspectivas, la del participante y la del observador, se produce una fisura semántica entre el vocabulario de corte normativo de los lenguajes cotidianos y del discurso ético, mediante los que tiene lugar una interacción comunicativa entre la primera y la segunda persona, y los lenguajes científicos, de diseño nominalista

y especializados en enunciados descriptivos. Se abre así un espacio que separa la norma de los hechos. La distinción consiste en que mientras que los vocabularios de las ciencias empíricas participan a un tiempo de la perspectiva del observador y de la del partícipe, los vocabularios normativos participan tan solo del discurso del partícipe. Aquellos tratan de explicar y describir un mundo objetivo que se presupone idéntico para todos los interlocutores. Estos tratan de justificar racionalmente las normas, valores e ideas destinados a regular nuestra conducta moralmente. El discurso normativo y el descriptivo han de asumir diferentes competencias, propias del modo en que hacen uso de las dos perspectivas metodológicas. Del mismo modo que cada una de las perspectivas es irreductible a la otra, también cada uno de los discursos es irreductible al otro. El discurso normativo no es reductible al discurso explicativo y descriptivo porque este tiene pretensión de verdad respecto a un mundo objetivo, haciendo uso de ambas perspectivas epistémicas, mientras que el discurso normativo solo puede ejercerse en la perspectiva del partícipe, desde la primera y la segunda persona, que permiten el establecimiento de normas bajo asentimiento reflexivo y razonamiento público. La ausencia de elementos intrínsecamente normativos en el mundo objetivo impide que un discurso objetivante como es el de la ciencia pueda hacerse cargo de cuestiones justificativas (Korsgaard, 2000; Apel, 1985; Habermas, 2006; Cortina, 2012a; Evers, 2010; Gould, 1999), que solo pueden afrontarse desde la perspectiva del partícipe propia del discurso normativo. Por estas razones en la ciencia se hace necesaria una labor complementaria de las diferentes perspectivas epistémicas, lo cual permitirá alcanzar un adecuado conocimiento sobre el mundo (Habermas, 2006: 212). Este conocimiento será, no obstante, insuficiente para resolver el problema de la fundamentación, del que se ocupará el discurso normativo, que permite establecer normas morales mediante el consenso intersubjetivo de los afectados por la norma en comunidades de argumentación, desde la perspectiva del partícipe.

La propuesta habermasiana supone un punto de referencia ineludible para atender a diferentes cuestiones que afectan a nivel metodológico al ejercicio de la neuroética. El discurso descriptivo y explicativo propio de las ciencias empíricas y el discurso normativo propio de la ética o filosofía moral presentan perspectivas metodológicas diferentes, con diferentes objetos de estudio y con diferentes competencias. Estos discursos, así como sus respectivas perspectivas metodológicas, son irreductibles el uno al otro, aunque no por ello incompatibles para ser utilizados de forma complementaria. La irreductibilidad de uno de los discursos al otro hace ilegítimo y rechazable todo intento de reduccionismo

cientificista. La neuroética habrá de operar desde el discurso descriptivo para ocuparse del ámbito explicativo de la moral, mientras que precisará del discurso normativo para abordar la dimensión justificativa de esta. La imposibilidad de reducir uno de los discursos al otro implica necesariamente que una neuroética crítica y no reduccionista lleve a cabo un ejercicio complementario de las ciencias empíricas y de la filosofía moral, haciendo uso de una dualidad de perspectivas metodológicas.

Capítulo 9. De la *ethica cordis* a una neuroética de la razón cordial

A lo largo de toda la investigación hemos tenido ocasión de apreciar el estrecho vínculo que existe entre la razón, por un lado, y las emociones y capacidades estimativas por otro. Este vínculo es revelado por multitud de estudios científicos y neurocientíficos, sin embargo su influencia no atañe solo a la dimensión natural de la moralidad. Un análisis filosófico de las condiciones de toda argumentación posible y válida ha de contemplar los aspectos cordiales implicados en el uso de la razón, que siempre se ejercita de forma encarnada. De lo contrario, caeremos en un concepto de razón pura que deja de lado toda una dimensión de elementos implicados en el diálogo, sin los cuales este no puede producirse. Tal es la propuesta de una *ethica cordis* o ética de la razón cordial, que actualiza las clásicas propuestas de la ética dialógica, deficitarias en la medida en que son excesivamente racionalistas. La obra de Adela Cortina transita desde las cuestiones más fundamentales de la ética, presentes a lo largo de toda su producción, hasta las cuestiones más actuales y problemáticas en el ámbito de la neuroética. Por ello, atenderemos a su propuesta como hilo conductor para transitar, también nosotros, desde una propuesta dialógica en clave cordial, una *ethica cordis*, hacia una propuesta metodológica crítica, que integre los resultados cosechados a lo largo de la investigación bajo el prisma de esta ética de la razón cordial, llamaremos a esta propuesta de método *neuroética de la razón cordial*.

9.1. Consideraciones críticas desde la *ethica cordis*

Muchas de las cuestiones a las que hemos atendido a lo largo de la investigación son abordadas por Cortina, cuya propuesta de una ética de la razón cordial supone un punto de referencia cardinal para elaborar una propuesta metodológica para una neuroética crítica en clave cordial, a la que referiremos como neuroética de la razón cordial. Su propuesta es heredera de la Escuela de Frankfurt, por ello la conexión de esta con Apel y Habermas se produce de forma estrecha, si bien cada uno de ellos presenta matices y características propios. Cortina actualizará las propuestas de estos, excesivamente racionalistas, incluyendo la consideración de los factores emocionales y estimativos implicados en el diálogo. Cortina trata específicamente temas como el de la fundamentación, tendremos ocasión de atender a estas cuestiones comentando algunos capítulos de *Ética mínima, Introducción a la filosofía práctica* (2007), en *Ética sin moral* (2010) y en *Ética aplicada y democracia radical* (2001). También se ocupa de abordar el

problema del reduccionismo naturalista y científicista en el ámbito moral, que tan central ha sido en nuestra reflexión, del que se ocupa especialmente en *Neuroética y Neuropolítica, Sugerencias para la educación moral* (2012a), donde se atiende a ello a la luz de las cuestiones más actuales de neuroética. La dimensión estructuralmente emotiva y racional de la moral, cuya explicación y descripción corresponde a la ciencia, en especial a la neuroética en su vertiente naturalista, casa a la perfección con la propuesta de una *ethica cordis* o ética de la razón cordial, que perfila nuestra autora en diversas obras, siendo *Ética de la razón cordial, Educar en la ciudadanía en el siglo XXI* (2009) la que reúne específicamente esta propuesta. Todas estas cuestiones son retomadas en obras posteriores, como *Aporofobia, el rechazo al pobre* (Cortina, 2017), en la que podemos ver cómo la ética, la neuroética, la ciencia y la razón cordial se relacionan a la hora de abordar problemas éticos y políticos de nuestras sociedades. Nuestro propósito es configurar una metodología interdisciplinar que permita complementar el ejercicio de las neurociencias de la conducta moral y la filosofía moral desde una perspectiva dialógica, en lo que podemos bautizar como una *neuroética de la razón cordial*.

Las funciones de la ética y la centralidad del problema de la fundamentación.

En el Capítulo 3 de *Ética mínima*, Cortina nos expone algunas cuestiones básicas acerca del ámbito de la ética. Define la ética como un saber práctico acerca de cómo deberían ser las cosas, que se diferencia del saber científico, el cual versa sobre las cosas tal y como son, sobre lo necesario, lo que no puede ser de otra manera a como es. Ya Aristóteles (2002) intuyó que esta diferencia entre ética y ciencia situaba a cada uno de estos campos del saber en territorios metodológicamente distintos.

Cortina explica que la ética se constituye como un metalenguaje filosófico respecto al lenguaje moral (también: Apel, 1985; Habermas, 2010, 2006), no pretende aumentar el número de normas o prescripciones morales, sino indagar en los motivos que tenemos para seguir esas normas, la ética no se ocupa ya de preguntar qué hemos de hacer, sino por qué hemos de hacerlo. Esa pregunta encarna el mismo sentido que la pregunta normativa de Korsgaard (2000), y ahonda en las raíces de esta fundamentación desde una perspectiva dialógica, afín a la de Apel y Habermas, pero actualizándola en algunos aspectos que resultan cruciales para nuestra investigación.

Cortina defiende que, dado que no podemos encontrar en el mundo objetos que sean moralmente normativos por sí mismos, tendremos que preguntarnos por las razones que nos conducen a la aceptación de una norma. Uno de los objetivos de la ética, defiende, consiste en indagar acerca de cuál es el fundamento de las normas y los juicios morales, el objetivo de la ética, por tanto, no es proponer contenidos, sino distinguir cuándo un contenido deviene moral dada su forma o dado el procedimiento que conduce a su enunciación y establecimiento. Cortina caracteriza el método filosófico como discursivo y autorreferencial, siendo esencialmente conceptual y argumentativo, aspectos que le diferencian metodológicamente del quehacer científico (Cortina, 2010: 30). En diferentes ocasiones especifica cuáles son las funciones de la ética, siendo la de la fundamentación una de las primordiales. Estas funciones serían las siguientes: a) dilucidar en qué consiste lo moral, b) tratar de ofrecer criterios para una fundamentación de lo moral y c) proponer una aplicación de los criterios morales propuestos desde la ética en el mundo real. (Cortina, 2010; Cortina, 2009: 44)

Crítica al cientificismo.

Tras estas consideraciones, Cortina se adentra de lleno en la cuestión metodológica. Propone explícitamente, bajo dos apartados titulados con los rótulos “métodos inadecuados” y “métodos adecuados”. Asegura que la ética solo puede alcanzar la verdad de la forma moral mediante el sistema filosófico, usando métodos también filosóficos. De ahí que el primero de los métodos que denuncie como inadecuado sea precisamente el método científico, en tanto que no es filosófico. De entre los métodos inadecuados que cita, este es al que más interesa atender para nuestra investigación. Ya que la neuroética combina el método científico con el filosófico, hay que ser prudentes a la hora de aplicarlos para no conceder a uno competencias que correspondan al otro (Buchanan, 2009; Apel, 1985; Habermas, 2006; Aranguren, 1994; Gould, 1999; Hume 1988, 1953; Moore, 1983; Salles, 2013, 2016, 2018; Evers, 2010; Cartwright, 2010; Teehan y diCarlo, 2004; García-Marzá, 2016; Cortina, 2007, 2012a).

El método descriptivo-explicativo, aplicado a la moralidad, es propio de las ciencias sociales — psicología, y sociología morales, antropología social y cultural— que son ciencias empíricas. Evidentemente, su objeto material pueden ser los contenidos morales, pero no poseen instrumentos para justificar la forma de la moralidad, ya que no pueden trascender al ámbito de los hechos empíricos sin propasarse en sus atribuciones. (Cortina, 2007: 66)

La investigación a través de este método culminará en una desesperanza dado el relativismo de contenidos que aparenta haber en diferentes culturas y sociedades. La contingencia histórica y social de esos contenidos les hace diversos (Ayala y Cela Conde, 2001; Hauser, 2008), eso no significa que exista un relativismo en relación a la forma (Apel, 1985; Habermas, 2006). La aplicación de un método adecuado debería posibilitar el descubrimiento de formas universales.

En el apartado 2 de *Ética mínima* (2007), “El problema de la fundamentación”, en el capítulo 3 de *Ética sin moral* (2010), “Panorama contemporáneo de la fundamentación de lo moral”, y en el capítulo 2 de *Ética de la razón cordial* (2009), Cortina vuelve sobre este tema, para someter a crítica las pretensiones reduccionistas de la ciencia al intentar naturalizar la fundamentación de la moral en lo que supone una tentativa de colonizar las competencias justificativas de la filosofía moral. La fundamentación, sostiene Cortina, se ha de producir como respuesta a una pregunta: ¿por qué? Ante ciertas normas, ideales o valores cabe preguntarnos qué es lo que hace a estos elementos moralmente normativos, por qué son buenas normas, buenos ideales o buenos valores. La respuesta puede apuntar a diversos factores, emocionales, psíquicos, socioeconómicos, religiosos, etc., pero habrá de apuntar hacia ellos racionalmente, es decir, quien presente una pretensión de validez respecto a una norma habrá de justificar racionalmente tal pretensión (Korsgaard, 2000; Apel, 1985; Habermas, 2006, 2010). Esa justificación ha de producirse, expone en consonancia con Habermas y Apel, en un marco dialógico y comunicativo que presente ciertas condiciones, donde los interlocutores puedan asentir mediante consenso a aquellas normas que puedan justificarse racionalmente. Esto no significa que la razón sea el fundamento de la moral, pero permite una fundamentación en los términos procedimentales descritos, que hemos tratado ampliamente en los capítulos precedentes.

La crítica que Cortina vierte sobre el científicismo consiste en poner en tela de juicio las pretensiones de abordar la cuestión del fundamento desde las ciencias empíricas y el método naturalista. El científicismo pretende reducir todo tipo de conocimiento posible a aquel que se ajusta a los parámetros de la ciencia moderna, es decir, a aquel tipo de conocimiento que versa sobre realidades materiales, sobre objetos del mundo accesibles a la experiencia.

Frente a la tradición europea de la filosofía de la reflexión, iniciada por Kant, que pretende desentrañar las condiciones que hacen posible el conocimiento teórico y el saber práctico y, por

tanto, la unión y fundamentación de ambos, el científicismo renuncia a denominar «saber» a la reflexión práctica y reduce a «objetivismo» la objetividad del conocimiento.

Si únicamente puede considerarse «saber objetivo» al que tiene ante sí el objeto, no hay más objetividad que la del saber científico y de ahí que «científicismo» signifique —en palabras de Habermas— «la fe de la ciencia en sí misma, es decir la convicción de que no podemos entender ya la ciencia como una forma de conocimiento posible, sino identificar el conocimiento con la ciencia». (Cortina, 2007: 90)

La ciencia no debería tener el monopolio de todo conocimiento posible. Los enunciados científicos, así como los enunciados normativos, son establecidos en contextos comunicativos de diálogo que aspiran a ajustarse a unas condiciones cuyo modelo es la idea regulativa de la situación ideal de habla (Apel, 1985; Habermas, 2010; Cortina, 2007, 2009). Tal y como hemos visto con Habermas en el capítulo anterior, los enunciados científicos tienen pretensión de verdad y los normativos pretensión de validez. Ambos tienen entre sus presupuestos la existencia de un mundo objetivo compartido por los interlocutores, idéntico para todos ellos. La ciencia reúne dos perspectivas al proponer enunciados con pretensiones de verdad, una de ellas es la perspectiva del observador, que trata de abordar en tercera persona la materia del conocimiento, la otra es la perspectiva del partícipe, en tanto que defiende sus enunciados con pretensión de verdad en contextos comunicativos donde interviene la primera y la segunda persona. Respecto a los enunciados con pretensión de validez, estos también son defendidos mediante argumentación en esos contextos, pero tan solo se adopta la perspectiva del partícipe, puesto que tales enunciados no refieren al mundo objetivo a la hora de establecer lo normativo. En cualquier caso, resulta posible un distanciamiento o descentramiento del sujeto respecto a los enunciados normativos, para su evaluación racional (Apel, 1985: 365-366; Habermas, 2006: 59). Así, ciencia y filosofía moral se constituyen como dos formas de saber diferentes, una de ellas refiere a un mundo objetivo compartido, la otra está dirigida a regular nuestra acción en él. Las pretensiones que defienden que puede objetivarse también la labor justificativa en torno a lo normativo cometen el error de creer que el tipo de conocimiento posible acerca de lo normativo puede participar de la perspectiva del observador. Esto conduce a un científicismo que es fuente de muchos errores en la investigación acerca de lo normativo y en los diferentes intentos de fundamentación que suscriben este científicismo (Spencer, 1964, 1883; Richards, 1986; Arnhart, 1998; Singer, 1981; Casebeer, 2003, 34; Flanagan; Cartwright, 2000: 328;

McShea y McShea, 1990, 1999; Churchland, 2013; Curry, 2006: 235, 237-238; Mora, 2007; Gazzaniga, 2006; Greene, 2013). La anterior cita de Cortina prosigue, haciéndose eco de estas cuestiones.

Ello [*considerar la ciencia como único conocimiento posible*] comportará la escisión teoría-praxis, conocimiento-decisión, reservando para la teoría y el conocimiento (científicos) toda posible objetividad y racionalidad, mientras que las decisiones morales quedan relegadas al ámbito subjetivo de los sentimientos y las preferencias irracionales. (Cortina, 2007: 90)

Tal y como hemos indicado, el error radica en considerar que solo desde la perspectiva objetivante de la ciencia puede obtenerse conocimiento. Precisamente la perspectiva del partícipe en los términos en que Habermas (2006), y también Evers (2010: 71), la describe es necesaria tanto para ejercer una tarea científica como una tarea filosófica normativa. Los enunciados establecidos mediante actos de comunicación, tanto si son enunciados con pretensiones de verdad como si lo son con pretensiones de validez, serán siempre susceptibles de revisión y corrección. De ahí que ninguno de los tipos de conocimiento pueda conquistar conocimiento más que en un sentido intersubjetivamente acordado en un contexto dialógico. La acusación de irracionalidad no puede verterse entonces sobre las decisiones morales tomadas mediante el procedimiento dialógico sin estar acusando también a la ciencia, dado que ésta también participa en este procedimiento. En capítulos anteriores hemos podido ver cómo Apel refiere a este problema en relación con la escuela popperiana y el Racionalismo Crítico (Popper, 1962, 1967, 1992). Lo cierto es que la acusación está infundada, precisamente este procedimiento es un modo de ejercer una reflexión racional en la esfera pública, siendo inclusiva, de modo que los resultados del procedimiento, sean enunciados con pretensión de verdad o con pretensión de validez normativa, sean aquello que se considera intersubjetivamente como más racional y mejor argumentado. El científicismo descrito presenta un sistema de complementariedad entre teoría y praxis, entre conocimiento y decisión, entre esfera pública y privada, del que se ocupa también Apel en *La transformación de la filosofía* (1985), como tuvimos ocasión de ver. Esta separación dibuja la idea de un conocimiento objetivo y neutral, por un lado, y una decisión subjetiva, comprometida y no neutral, sometida a voluntad, por otro. Así, el científicismo hace una inadecuada interpretación facticista del conocimiento y decisionista del valor (Cortina, 2010: 99). La escuela de Frankfurt denunció los inconvenientes de este científicismo (Apel, 1985), que Cortina recoge, y en los que nosotros centramos nuestra atención ahora.

Se suma a este esfuerzo las críticas de la Pragmática Trascendental (Habermas, 2010) y el Reconstructivismo de Erlangen (Lorenzen, 1969), que en esencia son las mismas.

La primera de las críticas acusa a la ciencia de ideología, en la medida en que carezca de un fundamento reflexivo. En tanto que la ciencia ignore los intereses que han movido a la razón a producirla, se desconocerá a sí misma (Habermas, 1999). Esto produce una visión deformada y deformante de la realidad, en sentido ideológico, como resultado de la errónea creencia de que la ciencia está desprovista de todo interés. En tanto que se considera a sí misma como desinteresada, sin estarlo, se convierte en buena candidata para producir un conocimiento contaminado por la ideología. Un interés que se desconoce dominará el pensamiento y la acción de forma inconsciente.

La segunda de las críticas es que el cientificismo bloquea toda fundamentación de la moral. La posibilidad de este bloqueo descansa en terreno poco firme, pues la propia ciencia necesita partir de ciertos presupuestos morales y asentarse sobre ellos (Apel, 1985: 377).

Efectivamente, la posibilidad de una objetividad científica exenta de valoraciones, no sólo no excluye, sino que presupone la validez intersubjetiva de normas éticas: no tiene sentido exigir la neutralidad axiológica de la ciencia empírica, en nombre de la objetividad, sin presuponer que la objetividad debe alcanzarse. (Cortina, 2007: 94)

Esta misma idea compone uno de los focos de la crítica que Apel dirige contra la ciencia, en tanto que pretende constituirse como un saber objetivo y no valorativo, mientras que en realidad en toda ciencia se haya presupuesta una ética como condición de posibilidad pragmático-trascendental (Apel, 1985: 377).

Una tercera crítica apunta la imposibilidad de argumentación sobre fines últimos si atendemos a los supuestos del cientificismo, que, además, se deciden dogmáticamente y no racionalmente. Ello supondría que la universalidad y la incondicionalidad del lenguaje moral, en tanto que fines últimos, serían inargumentables. Cortina apunta una cuarta crítica, en la que no profundizaré porque se aleja un tanto de la cuestión del cientificismo, esta crítica alude a la vinculación de la tesis de la complementariedad con las pretensiones de neutralización axiológica de la vida pública que encarna el liberalismo tardío, mientras que se reconoce el valor de una fundamentación privada de valores asociados a máximos de felicidad.

Cortina somete a crítica otros tres métodos a los que califica de inadecuados, atenderemos a estas críticas brevemente ya que no resultan tan relevantes para nuestra investigación como las relativas al método descriptivo-explicativo de la ciencia. El segundo de los métodos sometidos a crítica es el usado por los historicistas de la moral. El descubrimiento del origen histórico de los conceptos morales permite una explicación y comprensión de cómo estos surgieron y por qué, pero no pueden proporcionar criterios de justificación en el sentido normativo, ni tampoco criterios de verdad. El tercer método inadecuado es aquel que concibe la misión de la ética como recomendación de contenidos morales. El lenguaje de la ética no es prescriptivo, sino normativo. Por último, Cortina asegura que el análisis del lenguaje tampoco es un buen método, un método basado en describir proposiciones no podrá ocuparse adecuadamente de caracterizar lo normativo, se precisa una aclaración de las expresiones en su contexto pragmático, de ahí que la metaética —como ya apuntaba Apel (1985: 375)— no pueda ser normativa.

La ética dialógica y el problema de la fundamentación.

A pesar de las críticas al científicismo aún no hemos aclarado si una fundamentación ética es posible. Cortina dirige una crítica hacia el Racionalismo Crítico iniciado por Popper (1962, 1967, 1992) —como también hizo Apel—, en tanto que este racionalismo defiende una imposibilidad de fundamentación. El Racionalismo Crítico sitúa como blanco de sus objeciones el modelo clásico de racionalidad que parte del principio metodológico de razón suficiente, que presenta una combinación de verdad, certeza y fundamento. Pone así sobre la mesa algunos problemas relativos a la fundamentación de todo conocimiento posible. El primero de los problemas concierne al ámbito de la lógica formal, y trata de analizar algunas cuestiones problemáticas en relación a la validez de un argumento cuando tratamos el problema de la fundamentación. Tres son los datos que el análisis de esta cuestión nos ofrece. El primero es que a través del análisis lógico no se gana en contenido. El segundo es que la validez deductiva de un argumento no dice nada sobre la verdad de sus componentes —premisas o conclusiones—, solo garantiza que el tránsito del valor de verdad se deduzca de las premisas a la conclusión bajo el supuesto de que aquellas fueran todas verdaderas. Por último, un argumento deductivo inválido conforma una falacia. Cortina centra su atención en el segundo de estos datos, pues si la verdad de la conclusión no se transmite a las premisas eso significa que el fundamento no es más verdadero que lo fundado.

A esta cuestión hay que añadir otra que agudiza aún más el problema de la fundamentación. El universalismo del Principio de Razón Suficiente traslada al problema de la fundamentación el llamado trilema de Münchhausen. Este trilema plantea lo siguiente:

- 1) un regreso *ad infinitum* en busca de fundamentos, tras obtener respuesta a cada pregunta por el fundamento también puede exigirse fundamentación para dicha respuesta.
- 2) un círculo vicioso lógico en la deducción, en el cual se toman una y otra vez enunciados que precisaron fundamentación y que al ser falibles no conducen a un fundamento.
- 3) interrupción arbitraria del proceso en un punto determinado, estableciendo así un fundamento de forma dogmática.

El más común de los caminos es este último, un enunciado es tomado como autoevidente o autofundamentado, en virtud de lo cual se le inmuniza y se le convierte en dogma. En ese sentido el cognoscente opta por un *decisionismo* con el fin de alcanzar una fundamentación que de otra forma parece imposible.

Estos problemas derivan del propio planteamiento de la cuestión del fundamento, en tanto que se atiende al *Principio de Razón Suficiente*. Los problemas indicados se evitan si sustituimos este principio por el *Principio de Prueba Crítica*, según el cual se han de cuestionar bajo análisis racional todos los enunciados.

Puesto que el dogmatismo excluye la discusión racional, es necesario sustituir metodológicamente el Principio de Razón Suficiente por el de Prueba Crítica, que consiste en cuestionar todos los enunciados con la ayuda de argumentos racionales. Ello impide el nacimiento del trilema, porque afirmamos que no hay un punto arquimédico del conocimiento, sino que el hombre lo produce. (Cortina, 2007: 98)

En lo que concierne a la ética, la búsqueda de fundamentación conduce también al trilema, la interrupción arbitraria puede producirse de diferentes maneras:

- 1) tomando la intuición como fundamento, lo cual supone el problema de justificar las diferencias en las valoraciones.
- 2) axiomatizando, lo cual lleva a una regresión lógica *ad infinitum*.

- 3) apelando a la experiencia, con lo cual se introduce el problema de la inducción a la hora de justificar las diferencias entre valoraciones.

El Principio de Prueba Crítica también resuelve en el caso de la ética los problemas que supone el trilema, puesto que este evita al eliminar el estatus de postulado del Principio de Razón Suficiente, que pasa a considerarse una hipótesis más que debe contrastarse con otras. Sin embargo, atendiendo a las consideraciones que Apel aporta en torno a esta cuestión, el Principio de Prueba Crítica es problemático en ciertos aspectos, se ha de admitir que este Principio reconoce ciertos presupuestos pragmático-trascendentales *a priori* como condición de posibilidad de toda argumentación válida.

Teniendo en cuenta la crítica que Cortina vierte sobre el Principio de Razón Suficiente, por un lado, y las consideraciones en torno al Principio de Prueba Crítica, por otro, junto a la crítica de los métodos inadecuados, Cortina indica qué métodos resultan metodológicamente adecuados. Entre ellos sugiere la ética dialógica y la lógica trascendental. Ambos son métodos trascendentales, ambos pretenden una justificación racional, y ambos hacen uso del Principio de Prueba Crítica. La propuesta de la ética dialógica, además, parte del *factum rationis* de la argumentación. Es una alternativa mejor que la propuesta por la lógica trascendental, dado que la ética dialógica pone de relieve que esa lógica no es la única condición de posibilidad de este hecho, se precisa lo que Pierce (1965) denomina Semántica Trascendental, referido como Pragmática Universal por Habermas (2010) y Pragmática Trascendental por Apel (1985). La crítica de Apel contra el falibilismo va en esta línea, puesto que la discusión en torno al criterio falibilista precisa, como condición trascendental de posibilidad, unos presupuestos pragmáticos, relativos a esa dimensión pragmático-trascendental o semántico-trascendental. (Cortina, 2010: 105)

Las éticas dialógicas incluyen una dimensión social, de la que carecía la clásica propuesta kantiana, que permite que la reflexión se produzca de forma pública y el asentimiento ante la norma se produzca como resultado de la intersubjetividad, evitando así los problemas de un uso monológico de la razón. La cuestión de la fundamentación no consistirá, entonces, en buscar unos principios últimos de los que derivar lo fundamentado, sino que versará sobre las condiciones trascendentales de validez intersubjetiva de la argumentación.

Desde el punto de vista filosófico, la fundamentación de argumentos no consiste en decidirse por unos principios indemostrables, a partir de los cuales el argumento es derivable, sino en descubrir aquellos presupuestos sin los que la argumentación es imposible. El único procedimiento para hallarlos es la autorreflexión. (Cortina, 2007: 100)

Actualmente el problema de la fundamentación no se refiere a la cuestión moderna del origen de los conocimientos, sino a las condiciones trascendentales de la validez intersubjetiva de la argumentación, que pueden hallarse en la lógica (Kant), el sistema coherencial (Hegel), la semántica (Peirce) o la pragmática (Apel-Habermas). (Cortina, 2007: 101)

Una de las tareas más urgentes para la filosofía, por tanto, tendrá una función liberadora, o emancipatoria, como ya indicaba Apel, y consistirá en ofrecer una metodología crítica que permita eludir la aceptación dogmática de principios morales y ofrezca criterios de racionalidad crítica en lo concerniente a la fundamentación de normas morales, y, en tanto que comunidad de comunicación, también criterios metodológicos apropiados para la comunidad científica en su búsqueda del conocimiento. Todo ello, como puede verse, resulta especialmente relevante para la elaboración de una metodología apropiada para una neuroética crítica.

Y, porque no se diga que entender la filosofía como «crítica» y «liberadora» no es sino recurrir a palabras vacías, propongo a la ética una *tarea concreta*: la de *describir* su objeto con ayuda de las ciencias y del análisis lingüístico, concebirlo (es decir, expresarlo en conceptos), y *dar razón* de él (fundamentarlo), con el fin de obtener un canon *crítico* que haga posible la *argumentación*, reduciendo con ello en el ámbito práctico el dominio de los dogmas al lugar que tengan sentido. *Concepto, fundamentación racional y argumentación* permiten establecer un tipo de intersubjetividad universal, que es el mejor antídoto contra el dogmatismo. (Cortina, 2010: 34)

Este fragmento es de la mayor importancia. En él Cortina especifica la conveniente colaboración entre ciencias y filosofía a la hora de establecer descriptivamente el objeto de la ética, manteniendo la tarea de fundamentación en el ámbito filosófico, todo ello en un marco de argumentación, compartido entre ciencias y filosofía moral, aunque con especificidades metodológicas en relación a cada una de ellas, lo cual permitirá una intersubjetividad universal con una función crítica y antidogmática. Entendemos por dogma aquellas aserciones o prescripciones que se inmunizan ante la crítica racional apelando a “la *autoridad*, la *evidencia*, la *conexión inmediata con los sentimientos y las costumbres* o su *carácter metafórico*” (Cortina, 2010: 34; Albert, 1978). Es común en la ciencia, y concretamente en las neurociencias, apelar a algunos de estos criterios de forma dogmática, especialmente en el caso de la evidencia y de la

conexión inmediata con los sentimientos y las costumbres (Spencer 1964, 1883; Richards, 1986; Arnhart 1998; Singer, 1981; Casebeer 2003, 34; Flanagan; Cartwright 2000, 328; McShea y McShea 1990, 1999; Churchland 2013; Curry 2006, 235, 237-238; Mora, 2007; Gazzaniga, 2006; Greene, 2013). Por ello, una exigencia que debemos atender es la necesidad de hacer una crítica de las aportaciones de la ciencia y las neurociencias, así como de las diferentes propuestas en neuroética, en la medida en que puedan resultar reduccionistas.

Una ética de la razón cordial.

Antes de abordar las aportaciones de Cortina en el campo de la neuroética es preciso ahondar en su propuesta de una *ethica cordis*, cuya comprensión permitirá perfilar las exigencias y necesidades metodológicas de una neuroética crítica en clave cordial.

Si bien la propuesta de una ética de la razón cordial es transversal en toda la obra de Cortina, esta cristaliza de forma explícita en *Ética de la razón cordial* (2009). En esta obra, expresa la autora, el proyecto iniciado años atrás con *Ética mínima* (2007) es retomado con el objetivo de hacer una propuesta que incluya y supere aquella obra paradigmática en el pensamiento de la autora, su propuesta tomará la forma de una *ethica cordis* o *ética de la razón cordial*, que está caracterizada por reconocer no solo la necesidad de la razón en el marco de las éticas dialógicas, sino también de las emociones, sentimientos y capacidades estimativas que permiten que el diálogo se produzca realmente, ya que disociar la razón del corazón no parece posible, pues todo argumento y toda discusión se da siempre de forma encarnada (Conill, 2006, 2015b, 2016). La razón es siempre una razón cordial, vinculada a una dimensión emotiva que no puede ser ignorada si queremos conocer las verdaderas condiciones en las que se produce el diálogo, tanto en el ámbito científico como en el normativo, y más específicamente en este último, ya que se produce una *ligatio*, emocionalmente motivada, como obligación moral. En este sentido, las propuestas de corte más racionalista, como las de Habermas y Apel, presentan deficiencias en lo concerniente a aspectos emotivos y estimativos vinculados a la razón y al diálogo, como el carácter, las virtudes, los valores, el sentir común, los sentimientos, etc. (Cortina, 2009: 25)

Esta es la tarea que se propone *Ethica cordis*: intentar superar las limitaciones de una ética mínima procedimental, actualizando a la vez sus planteamientos en una ética que no es sólo de la razón procedimental, sino de la razón humana íntegra, de la razón cordial (Cortina, 2009: 32)

Cortina se propone así perfilar una ética del discurso capaz de integrar los aspectos emocionales y estimativos implicados en toda argumentación, especialmente en la que tiene lugar en el discurso normativo.

“[...] una “versión cálida” de la ética del discurso, una versión preocupada no sólo por los aspectos *epistémicos* del vínculo comunicativo, sino también por los aspectos *cordiales*, que lo convierten en verdadera comunicación. (Cortina, 2009)

Uno de los primeros elementos de corte emocional que afectan a la razón en el nivel argumentativo es el de reconocimiento. El reconocimiento no se produce solo como una condición pragmático-trascendental en toda argumentación (Apel, 1985: 404; Habermas, 2006: 24-25), sino que es un reconocimiento que además se produce en términos emotivos (Colmenares, 2005; Morgado, 2010; Evers, 2010; Eisenberg, 2000; Tagney, 2007; Damasio, 2008; Waal, 1997, 2007; Brosnan y Waal, 2003; Tomasello y Vaish, 2012; Churchland, 2013), es un reconocimiento cordial, que tiene en su base no solo razones de peso, sino un componente sintiente y compasivo que sitúa el nivel de ligatio moral en el plano social (Piaget, 1987; Kohlberg, 1992; Gilligan, 2003). Nos sentimos moralmente obligados con los demás, y no es aquí tanto la razón, sino el corazón, el que motiva la acción moral y, como condición de posibilidad de la misma en un marco dialógico, el reconocimiento recíproco. (Cortina, 2009: 51).

El siguiente aspecto cordial que Cortina reconoce como necesario en la ética es el interés, que puede ser en algunos casos egoísta, basado en un cálculo inteligente y estratégico para la propia supervivencia, bienestar o beneficio (Waal, 1997, 2007; Brosnan y Waal, 2003; Evers, 2010; Churchland, 2013), pero también existe el interés basado en la apreciación de lo valioso, en el altruismo —en un sentido no solo biológico— (Ayala y Cela Conde, 2001), en la indignación ante lo injusto, en la estima por la propia vida y la de otros (Kant, 1992, 2000; Korsgaard, 2000), etc., elementos que además de encontrarse en el ámbito moral también están presentes en el político (Cortina, 2001, 2009; Krause, 2009; Marcus, 2002). En definitiva, existe un interés, emocionalmente mediado y en ocasiones también emocionalmente motivado, que es parte indispensable de todo razonamiento en el contexto de un diálogo real y encarnado. (Cortina, 2009: 57-77)

Para comprender esta razón cordial, Cortina atiende a las principales emociones que influyen a la razón en el ámbito práctico, las emociones sociales. Estas emociones

tienen una base natural, pero pueden cultivarse cultural y educativamente (Evers, 2010; Churchland, 2013). Entre ellas destaca el sentimiento de simpatía, gracias al cual, dice Cortina, es posible “transitar de la obligación natural a la obligación moral”, añadiendo que “son los sentimientos sociales los que mantienen moralmente a la sociedad” (Cortina, 2009: 84; Cortina, 2017). Como hemos visto en capítulos anteriores, la razón por sí sola es insuficiente para la formación del agente moral, este precisa un nivel de prosocialidad que dirija sus consideraciones morales a los demás, y no a su propio interés de forma egocéntrica. Un cálculo puramente racional sobre lo moralmente correcto estaría dejando fuera de consideración aspectos que son constitutivos de la moralidad humana. El caso de la psicopatía, en el que las facultades emocionales y las emociones sociales no presentan una funcionalidad adecuada, es una muestra de cómo una moralidad excesivamente racionalista y egoísta hace que difícilmente podamos considerar al agente como moral en el sentido más genuino del término (Darwin, 1979; Haidt, 2001: 816; Damasio; Greene; Raine, 2000; Blair). Es preciso incluir y atender a la esfera emocional y estimativa para que la deliberación sobre lo moral se produzca de forma adecuada, se precisa, en definitiva, una razón cordial, encarnada. De forma muy cercana al sentimiento de simpatía encontramos la compasión, que sirve como motivo de la acción dirigida a paliar el sufrimiento o el estado desfavorable de los otros (Waal, 1997, 2007).

En lo concerniente a estas emociones, somos, al menos en parte, moralmente responsables de las mismas, ya que pueden cultivarse y podemos aprender a gestionarlas y controlarlas, por eso los conceptos de agencia emocional (Sherman, 1999) y de inteligencia emocional (Goleman, 1996) se vuelven aquí centrales. No somos, al menos no necesariamente, seres afectados de forma pasiva por las emociones, siendo estas las que nos asaltan, al contrario, podemos dirigir nuestras capacidades emocionales y responsabilizarnos de las mismas, podemos cultivar un carácter, el modo en el que sentimos, y cómo esa dimensión emocional es gestionada junto a la racional en el ejercicio de nuestras conductas morales y éticas.

La reputación es otro de los elementos que encontramos en el nivel emocional. La simpatía no solo se da desde el sujeto hacia los demás, sino también en la dirección opuesta, desde el grupo o la sociedad hacia cada uno de sus miembros. Las pasiones del orgullo y de la humillación, así como la posibilidad de ser reconocido y estimado por el grupo o de ser rechazado y odiado, componen uno de los engranajes sociales más

arraigados en nuestras emociones y en nuestra naturaleza, aunque pueda ser cultivado, modulado y modificado cultural y educativamente (Evers, 2010; Churchland, 2013; Hauser, 2008). Contar con el beneplácito del grupo, el afán de reconocimiento y evitar el rechazo o simplemente la no consideración por los demás, es uno de los ejes de la vida social en general (Colmenares, 2005; Waal, 1997, 2007; Tomasello y Vaish, 2013), y de la moral en particular (Kohlberg, 1992; Piaget, 1987). Si bien la reputación presenta un origen, desarrollo y base biológica, se halla además presente en niveles más teóricos de reflexión, como el del reconocimiento en toda argumentación posible (Apel, 1985: 404; Habermas, 2006: 24-25; Cortina, 2009: 51; Honneth, 1997).

Frente a la mano de hierro del Estado y a la mano invisible del mercado, Cortina plantea la mano intangible de la sociedad civil (Brennan y Pettit, 2004) como una economía de la estima, que sugiere llamar *kudonomía*, basada en ese afán de reputación. Según esta *kudonomía*, los miembros de las sociedades serían premiados y sancionados por estas otorgando y privando de reconocimiento y estima a sus miembros, según sus méritos sociales y morales. A pesar de las dificultades que esta propuesta presenta, y que la propia Cortina señala, resulta especialmente interesante cómo elementos emocionales como el sentimiento de simpatía, la estima y la reputación pueden regular el funcionamiento de las sociedades y moldear el carácter y las acciones de los miembros que la forman, tanto a nivel social como moral.

Ahora bien, estos sentimientos o estas emociones no son suficientes para explicar las fuentes de la obligación moral. La intervención de la razón es necesaria para ello. Del mismo modo, aquellas explicaciones basadas únicamente en el uso de la razón serán insuficientes. Es preciso que ambas esferas, la emocional y la racional, sean ejercidas conjuntamente en la labor de investigación y búsqueda de fundamentos en torno a la obligación moral y las fuentes de la normatividad.

Esta línea de investigación conduce a Cortina a una reflexión sobre las condiciones trascendentales de la moralidad teniendo en cuenta no solo los aspectos racionales, sino también los emocionales y estimativos. Profundiza en este punto en el concepto de autonomía. La libertad y la capacidad de autolegislación moral de la voluntad son una fuente de obligación moral en todo ser racional, y no solamente desde un sujeto hacia sí mismo, también puede producirse el ejercicio de la autonomía intersubjetivamente en una comunidad de comunicación. Esa autolegislación se produce

desde el ámbito de la razón práctica. En ese sentido, la obligación moral, en la medida en que es racional, es universalizable, todo ser dotado de razón debería verse obligado moralmente bajo ciertas condiciones de argumentación racional. Hasta aquí la propuesta de Cortina es totalmente afín tanto a la propuesta kantiana (Kant, 1992, 2000) como a la de sus herederos pertenecientes a la corriente dialógica (Apel, 1985; Habermas, 2010), en esta última el carácter monológico del uso práctico de la razón, que planteaba Kant, se traslada al marco dialógico de una comunidad de comunicación, conservando sus principales características, que son la autonomía, la argumentación racional, la prevalencia de los mejores argumentos y la pretensión de universalidad. En las éticas dialógicas la autonomía es la de una comunidad de hablantes, en la que se busca el mejor argumento aportado por los afectados de una norma, generando así un consenso en condiciones de racionalidad, que debería obligar universalmente. Así, Cortina se inscribe en la tradición dialógica procedimental, heredera de las éticas formales kantianas. Sin embargo, denuncia que el procedimentalismo puramente racional propio de esta tradición (Apel; Habermas) es insuficiente para dar cuenta de las condiciones en las que se produce el diálogo, tanto en la comunidad ideal de habla como, y especialmente, en las comunidades reales de argumentación, ya que la dimensión cordial ha de ser tomada en cuenta, en tanto que constitutiva de todo sujeto, que no argumenta desde una razón pura, sino desde una razón encarnada (Conill, 2006), en la que están presentes diversos aspectos cordiales.

Entre estos elementos emocionales el sentimiento de respeto, estrechamente vinculado al reconocimiento de la dignidad, es decir, el reconocimiento de que hay seres valiosos por sí mismos y que no tienen precio, es el principal de los móviles para la acción moral. Este móvil es cordial y está emotivamente constituido, aunque pueda ser racionalmente reforzado. Es decir, encontramos que los motivos para la acción moral, aunque influidos por la razón, son esencialmente cordiales. Y es en ese punto en el cual la *ligatio* con los demás transita hacia la obligación moral, nos *sentimos* moralmente obligados hacia los demás. A la razón corresponde la tarea crítica de fundamentar y justificar esa obligación cuando sea posible, desechándola cuando no resista la crítica (Haidt, 2001, 2013).

Para que sea posible este reconocimiento mutuo y recíproco de la dignidad, así como algunos logros sociales y morales como los derechos humanos, que dependen del

reconocimiento de la dignidad, es preciso la capacidad estimativa, la capacidad de valorar, y de hacerlo en un sentido más amplio que el meramente biológico (Churchland, 2013), aunque intervengan en esas estimaciones aspectos de una estructura biológica que está a la base. Para cultivar la degustación de valores, el reconocimiento creativo de los mismos y el progreso de las sociedades en concordancia con estos valores, son necesarias unas capacidades que no dependen solo del puro ejercicio de la razón, sino también de un nivel emotivo y sintiente que permite que la estimativa no quede relegada al frío y teórico ámbito de un discurso puramente lógico y racional, sino que se produzca de forma encarnada y cordial, que es el modo en que tanto la reflexión como la valoración moral pueden producirse de forma plena y real. Resultan relevantes en este punto el conjunto de elementos necesarios para la agencia moral que estudiamos en el Capítulo 2, donde las facultades sociales y emocionales (teoría de la mente, empatía, etc.) así como las emociones, especialmente las sociales, resultaban necesarias para constituir al agente moral (Darwin, 1979; Ayala y Cela Conde, 2001; Ayala, 2001; Waal, 1997, 2007; Tomasello y Vaish, 2013; Kohlberg, 1992; Piaget, 1987).

Así, Cortina aporta algunos elementos cordiales que han de ser considerados en el ejercicio de la autonomía y de la razón, tales como la conciencia moral (Cortina, 2016, 2017), que está impregnada de elementos emocionales aunque pueda ser mediada por la razón (Darwin, 1979; Kant, 1992), el sentimiento moral, entendido como capacidad de sentir placer o agrado al realizar acciones acordes con el deber y rechazo o dolor con las contrarias al deber, la autoestima o respeto por sí mismo, el amor al prójimo, etc. Estos elementos, dice Cortina, son “condiciones *subjetivas* de la receptividad para el concepto del deber, no condiciones objetivas de la moralidad” (Cortina, 2009: 120). Es decir, entre las condiciones de posibilidad de la moralidad encontramos elementos emocionales, sin los cuales la moralidad no sería posible, o al menos no tal y como la conocemos. La moral es constitutivamente racional, pero también constitutivamente cordial.

Cortina conduce todas estas consideraciones hacia una propuesta dialógica cordial, en la que el reconocimiento, la autonomía, el establecimiento de normas, la pretensión de universalidad, etc., se produzcan en una comunidad de comunicación y de argumentación, que deberá atender al ejercicio de la razón, pero al mismo tiempo no podrá desestimar el papel imprescindible de las emociones, los sentimientos y las capacidades estimativas.

Es preciso contar con razones de la razón y con razones del corazón para hablar de justicia, no se puede reducir la comunicación al discurso lógico. (Cortina, 2009: 161)

Cortina, como Apel y Habermas, considera que tanto el discurso científico como el discurso normativo se producen en el seno de comunidades de comunicación, variando únicamente en el tipo de enunciados que emiten y el objeto de estos, a saber, enunciados con pretensión de verdad dirigidos a un mundo externo, y enunciados con pretensión de validez, dirigidos a regular la acción en ese mundo, respectivamente. Este tipo de argumentación “recibe el nombre de ‘discurso’, que será teórico cuando se trate de la verdad; y práctico, cuando se trate de la corrección o justicia de las normas” (Cortina, 2009: 179).

Como Apel y Habermas, también Cortina atiende a las condiciones de posibilidad necesarias para que todo discurso, teórico o práctico, pueda producirse. Entre esas condiciones reconoce también nuestra autora la aceptación de una norma ética fundamental, sin la cual el discurso no es posible, que consiste en el reconocimiento de otros como interlocutores válidos, en plano de igualdad, atendiendo a la fuerza del mejor argumento. Este reconocimiento se configura como un *reconocimiento cordial* o *reconocimiento compasivo*, basado en la compasión como el sentimiento que conduce a preocuparnos por la justicia (Cortina, 2009: 215). Reconoce también la autora que, aunque no tengamos garantías de alcanzar el ideal de comunidad de comunicación, debemos atender a él como una idea regulativa. La aportación específica de la autora, en este punto, consiste en denunciar que el ejercicio del diálogo precisa de elementos que Apel y Habermas no contemplan en sus propuestas, elementos cordiales, como lo son sentirse atraído por determinados valores o incorporar cierto carácter moral o ethos, es decir, es preciso atender a un concepto razón cordial. En función de este concepto de razón el reconocimiento, del que ya hablan Apel y Habermas, se configura como un reconocimiento cordial.

[...] para argumentar con éxito sobre lo justo [la ética] ha de hundir sus raíces en su vertiente cordial y compasiva. La razón íntegra es entonces razón cordial, porque conocemos la verdad y la justicia no sólo por la argumentación, sino también por el corazón. (Cortina, 2009: 191)

Cortina especifica además cómo sería un carácter adecuado para la comunicación. Toma de Pierce (1965) cuatro aspectos que este propone en relación a la comunidad científica, pero que son necesarios, además, para un carácter predispuesto a la

comunicación, sea cual sea el tipo de discurso al que refiramos, siempre que sea comunicativo. Estos aspectos son:

- a) la apertura a la crítica racional, que permita someter a revisión nuestras creencias y aserciones.
- b) el reconocimiento, que implica necesariamente reconocer los derechos del resto de interlocutores y la garantía de que sus argumentos sean tomados en consideración en plano de igualdad.
- c) el compromiso, que en el caso del discurso moral es compromiso con la justicia, mientras que en el discurso científico lo es con la verdad.
- d) la esperanza de que será posible llegar a un consenso, sobre intereses universalizables en el discurso práctico y sobre la verdad en el discurso teórico.

En el ámbito social plantea Cortina los aspectos básicos de una ética cívica cordial, que sería la cristalización de la ética de la razón cordial en el ámbito de la sociedad. Los principios de esta ética cívica cordial serían:

- a) no instrumentalización de las personas.
- b) empoderamiento o principio de las capacidades.
- c) principio de justicia distributiva, basado en reparto equitativo de cargas y beneficios.
- d) principio dialógico, en función del cual todos los afectados por las normas han de ser tenidos en cuenta en la discusión, ya sea como interlocutores reales o virtuales.
- e) principio de responsabilidad por los seres indefensos no humanos.

La propuesta de Cortina permite trasladar la cuestión de la fundamentación ética al problema de fundamentar ciertos derechos previos a todo pacto o consenso, dado que estos derechos encarnan condiciones necesarias de deliberación que permiten el procedimiento mediante el cual el pacto o consenso puede darse.

Según su propuesta, como ya hemos visto, son dignos y susceptibles portadores de derechos todos aquellos seres con capacidad comunicativa, sea real o virtual. Por ello, para que el discurso sea justo debe contar con una serie de derechos entre sus condiciones de posibilidad. Concernientes a la realización del propio discurso, se ha de reconocer el

derecho de participación, que implica derecho a poner en cuestión cualquier afirmación, y a expresar la propia posición, necesidades e intereses, y derecho a no ser obligado y a no decidir por coacción. Habría que mencionar aquí también el derecho a la vida, sin vida no hay diálogo posible. Un segundo tipo de derechos tomarían la forma de libertades, estos pueden identificarse con algunos derechos humanos, y son precisos para que la norma resultante del diálogo tenga validez. Son la libertad de conciencia, la libertad de expresión y de información, la libertad religiosa y de opinión, la libertad de asociación y el derecho a participar en la deliberación sobre normas que me afectan. Se precisan además unas condiciones materiales y culturales que permitan dialogar y decidir en pie de igualdad, los derechos que velan por garantizar el cumplimiento de estas condiciones son los que históricamente se han reconocido como derechos económicos, sociales y culturales. Todos ellos no responden solo al cálculo racional sobre lo normativo, sino que incorporan, como hemos visto, toda una dimensión cordial, en la cual están incluidos el reconocimiento, la estimativa y la compasión, tan necesarios para deliberar sobre la justicia.

9.2. Neuroética de la razón cordial: criterios metodológicos para una neuroética crítica e interdisciplinar

Como hemos podido ver, Cortina recoge los aspectos más relevantes de las propuestas dialógicas en lo concerniente al método en las ciencias empíricas y en la filosofía moral, actualizándolos con la propuesta de un concepto de razón como razón cordial. Todas estas consideraciones son tenidas en cuenta por Cortina a la hora de abordar la relación entre neurociencias y ética, tarea de la que se ocupa en diversas contribuciones, como libros y artículos científicos, entre los que cabe destacar *Neuroética y neuropolítica* (2012a). Si pretendemos proponer un método adecuado para una neuroética crítica en clave cordial, analizar las cuestiones que la autora trata en esta obra resulta crucial, nos permitirá tender un puente entre su propuesta de una neuroética de la razón cordial y nuestra propuesta de una metodología apropiada para la neuroética.

Neurociencia, ética y el problema del reduccionismo.

Los problemas de la fundamentación y del método que hemos podido ver atendiendo a *Ética mínima* son actualizados por Cortina a la luz de los problemas de cientificismo que se han articulado en relación a las neurociencias. En el ámbito de las

neurociencias estos problemas cristalizan en pretensiones de naturalizar la fundamentación de lo moral, lo cual implicaría que el método adecuado para abordar este problema sería el propio de las ciencias naturales y no el de la filosofía.

Si bien el contexto de aparición de estas pretensiones científicistas es nuevo, dada la proliferación de las neurociencias, el problema que supone este reduccionismo es el mismo que Cortina abordó en *Ética mínima* hace décadas, y que retoma recurrentemente a lo largo de su obra (Cortina, 2010: 98-119; Cortina, 2012a, 2012b). Además, el problema del reduccionismo naturalista se muestra como una constante a lo largo de la historia, en la cual a menudo surgen conatos de reduccionismo científicista en lo que concierne a la moral, asociados a diferentes proyectos y disciplinas científicas (Spencer, 1964, 1883; Richards, 1986; Arnhart 1998), incluyendo entre ellas las neurociencias (Casebeer, 2003, 34; Churchland, 2013; Mora, 2007; Gazzaniga, 2006; Greene, 2013).

Cortina reivindica el papel de la filosofía como una demarcación del problema al que ella refiere como “el pugilato natural-sobrenatural”. Frente al naturalismo por el que abogan algunos autores, suele situarse lo sobrenatural como alternativa. Pero no se toma en consideración la reflexión filosófica que, teniendo en cuenta los datos que la ciencia aporta sobre nuestro origen evolutivo y nuestra constitución biológica, ofrece una rica y necesaria reflexión moral que se demarca de la disyuntiva natural-sobrenatural. Precisamente reivindicar el papel de la filosofía a la hora de abordar cuestiones como la de la justificación o la fundamentación es uno de los objetivos de nuestra investigación. Formular un método en esta nueva disciplina que es la neuroética, o al menos algunas consideraciones que orienten la praxis metodológica, se presenta como una prioridad.

En *Neuroética y Neuropolítica* (2012a) Cortina actualiza sus principales críticas al científicismo a la luz de los problemas que este plantea en las neurociencias. Trata ampliamente la cuestión, para ello hace una exposición de qué es la neuroética y aborda problemas relativos a la fundamentación, a la cuestión de la libertad, al neuroderecho, a la neuropolítica y a la neuroeducación, principalmente. No atenderemos en este capítulo a la cuestión de la libertad, ya que nos ocupamos de ello en el segundo epígrafe del Capítulo 4 de la presente investigación.

El científicismo en neurociencias plantea que, si las neurociencias descubren unos códigos morales inscritos en el cerebro gracias a la evolución, podrá fundamentarse una

ética universal siguiendo esos códigos, ya que todos los seres humanos compartimos el órgano que nos hace morales, el cerebro. Algunos neuroeticistas, al hablar de fundamento, confunden un fundamento último de la normatividad con las estructuras psicobiológicas sobre las que la conducta moral se asienta. En relación a esta confusión Cortina distingue entre los conceptos de *base* y *fundamento* de lo moral. La autora reconoce que “las descripciones asépticas de lo moral serán cosa de las ciencias empíricas, pero no de la filosofía moral” (Cortina, 2010: 45), a esta última corresponderá una función normativa, estando entre sus competencias abordar el problema de la fundamentación.

Respecto al concepto de *base*, señala que el cerebro es condición necesaria para realizar tareas morales como sentirse obligado por normas morales, captar valores que denominamos morales, tener emociones y sentimientos de carácter moral, o desarrollar y ejercitar virtudes. El cerebro es el soporte físico que otorga las estructuras para la moral, pero eso no significa que las determine (Ayala y Cela Conde, 2001; Hauser, 2008). En esas estructuras han de asentarse una serie de contenidos morales, en este punto el cerebro es plástico y las posibilidades son múltiples (Evers, 2010, 2015; Aranguren, 1994; Jácomo y Serrão, 2013; Salles, 2017). Esta base refiere solo a las estructuras psicobiológicas de la conducta moral, que permiten describirla y explicarla, pero no justificarla en sentido normativo (Korsgaard, 2000; Buchanan, 2009; Apel, 1985; Habermas, 2006, 2010; Cortina, 2007, 2009; Aranguren, 1994; Teehan y diCarlo 2004; Gould, 1999; Hume, 1988, 1953; Moore, 1983; Salles, 2013, 2016, 2018; Evers, 2010). Dependerá de la filosofía moral, no de las neurociencias, el llevar a cabo una buena tarea de fundamentación a la hora de plantear una ética universal.

Si existiese un fundamento cerebral para una ética universal, eso querría decir que la evolución ha marcado en nuestro cerebro y en nuestros genes no solo la estructura de lo moral sino también los contenidos morales, que habrían de ser determinados por los neurocientíficos para después exportarlos al ámbito de la ética como normas morales de carácter universal. Las posturas que defienden esta idea protagonizan tentativas científicas, que no son sostenibles, como hemos tenido ocasión de analizar a lo largo de nuestra investigación.

Cortina indica que la pretensión de una universalidad moral en neurociencias puede distinguirse en dos posibles posturas. Una de ellas pretende una fundamentación ética que se corresponda con aquellos códigos para los que nos predispone nuestra base

moral. Esta postura desplaza a las propuestas anteriores, las neurociencias heredarían el lugar que antes ocupaban las teorías éticas y las diferentes propuestas morales dado que, como pretenden algunos autores (Cortina habla aquí de: Gazzaniga, 2006; y Mora; 2007; también podemos incluir a: Churchland, 2013; Greene, 2013; Casebeer, 2003), la neurociencia sería capaz de extraer del cerebro una ética universal en tanto que todos contamos con este órgano y con ciertas pautas de conducta predispuestas por él en función de nuestra herencia biológica. Estas posturas, por las razones expuestas en capítulos precedentes, incurren en reduccionismo naturalista que no resiste la crítica. Por otro lado, están aquellas posturas que diferencian entre estas predisposiciones biológicas, que constituyen un hecho natural, y aquellas normas que pretenden orientar la acción en sentido moral. Estas posturas, señala Cortina, son compatibles con las propuestas de la filosofía moral. No pretenden aportar unos contenidos éticos universales, sino más bien descubrir las estructuras morales que subyacen en todas las formas de vida humana y que tienen como soporte el cerebro. Las diferentes culturas modularían esas estructuras universales. Estas posturas recurren a métodos filosóficos en su tarea y son compatibles con propuestas éticas anteriores (Cortina habla aquí de: Levy, 2007; y Hauser, 2008; podemos incluir también: Evers, 2010). Estas últimas propuestas, que abogan por investigar la estructura moral y no la justificación de contenidos, sin reducir a las bases de la moralidad la cuestión del fundamento, dice Cortina, son más prometedoras. Y lo son en virtud de que resultan metodológicamente más adecuadas, como hemos podido analizar a lo largo de la investigación.

Cortina manifiesta una preocupación en relación al modo en que los avances de las neurociencias pueden ser asumidos y aplicados en el plano ético. Conceptos como el de justicia o el de bien moral, así como el modo en que los aplicamos moralmente, no son algo que venga determinado genéticamente ni cerebralmente. Si bien tenemos predisposiciones a la hora de llevar a cabo nuestra conducta moral, no debemos tomar estas como códigos morales inscritos por la naturaleza en nuestra biología, sino más bien como una base que nos condiciona, pero que es maleable y presenta múltiples posibilidades de realización (Evers, 2010; Hauser, 2008; Jácomo y Serrão, 2013; Salles, 2017).

Tras exponer detalladamente los principales avances en neurociencias en relación a la conducta moral, Cortina extrae algunas conclusiones, haciendo balance de estas

(Cortina, 2012a: 82-91). Atendiendo a ello, nuestra investigación ha tratado de llevar a término las recomendaciones sugeridas por Cortina, que pueden verse reflejadas en los siguientes puntos, siendo estos algunos de los ejes de nuestro estudio.

- 1) Hemos investigado a fondo la distinción entre bases y fundamentos, y sus implicaciones en el ámbito de la filosofía moral.
- 2) Hemos investigado cuáles son las estructuras de la moralidad en nuestra especie, atendiendo a su origen filogenético, su desarrollo ontogenético y su funcionamiento cerebral; asimismo hemos señalado el modo en que los diferentes contenidos, que son contingente e históricos, se relacionan con esas estructuras.
- 3) Hemos tratado de aclarar terminológicamente conceptos como los de naturalismo, naturaleza, naturaleza humana, naturaleza moral humana, así como su relación con lo moral y con la ética. También hemos aclarado en qué consiste el reduccionismo naturalista y los principales modos en que este se produce en el ámbito de las ciencias empíricas.
- 4) Hemos reconocido el papel de las emociones en la vida moral, analizando qué elementos constituyen al agente moral como tal, desde una perspectiva filogenética, ontogenética y funcional, y desde el estudio del juicio moral; analizando el modo en que se relaciona con el uso de la razón, así como los problemas metodológicos que encontramos en relación con ello; también hemos analizado el papel de las emociones y la capacidad estimativa en el ámbito moral, desde la perspectiva de una ética de la razón cordial.
- 5) Hemos estudiado diversas propuestas relativas a la formación y el procesamiento de los juicios morales, para aclarar el modo en que estos se producen, sometiendo a crítica aquellos aspectos que hemos considerado oportuno.
- 6) Hemos expuesto cómo las estructuras morales, formadas por ciertos elementos que son condición de posibilidad para la agencia moral, están relacionados con usos y funciones que no siempre son específicos de la esfera moral; hemos visto, además, que las áreas cerebrales implicadas en el uso de esos elementos no son específicas de la moralidad, sino útiles para muchas otras funciones.
- 7) Hemos llevado a cabo revisiones terminológicas en relación a muchos puntos, por ejemplo en relación al concepto de naturaleza, en relación a los factores evolutivos implicados en el origen evolutivo de la moralidad, en relación al juicio moral, y en relación a la cuestión del libre albedrío, entre otros.

- 8) Hemos señalado que el ser humano tiene una tendencia natural a actuar de forma egocéntrica y endogámica, indicando los modos en que esta tendencia se puede moldear y contrarrestar, siendo esta una de las funciones de la neuroeducación moral.
- 9) Hemos señalado que las predisposiciones conductuales heredadas biológicamente e inscritas en nuestra naturaleza no pueden servir para fundamentar ningún tipo de norma moral.
- 10) Hemos señalado cómo el desarrollo cerebral y la adquisición de contenidos permiten malear la conducta, contrarrestando las tendencias conductuales heredadas, e incluso cambiándolas a nivel genético, tal como vimos con la propuesta de una epigénesis proactiva; en todo caso las estructuras quedan siempre modificadas por los contenidos, y corresponderá a la filosofía moral hacerse cargo de justificar los valores, normas y principios en los que hemos de educar.

Con todo ello no pretendemos ofrecer una solución a cada uno de los problemas planteados. Lo que pretendemos es encaminar la investigación en la dirección adecuada, haciendo algunas aportaciones interesantes, tanto en el ámbito descriptivo y explicativo como en el justificativo. Pero sobre todo pretendemos descubrir los aspectos en los que las neurociencias y la filosofía moral pueden operar de forma conjunta siendo fructíferas en su labor compartida. De este modo podremos especificar qué aspectos son metodológicamente recomendables para una neuroética crítica, que elaboraremos en clave cordial.

Dimensión cordial en el método de una neuroética crítica.

Muchos de los aspectos metodológicos relevantes para el adecuado ejercicio de una neuroética crítica han sido atendidos a lo largo de la investigación, entre ellos encontramos la dimensión cordial de la razón. Nos ocuparemos, en este apartado, de especificar cómo se integra esta dimensión cordial en el método de una neuroética crítica.

La perspectiva excesivamente racionalista de las propuestas dialógicas que preceden a Cortina es deficitaria en la medida en que no atienden al papel de las emociones y las capacidades estimativas, de las que no podemos desvincularnos cuando hacemos uso de la razón. Un estudio completo de las condiciones de posibilidad de toda

argumentación posible y válida ha de incluir, por tanto, la dimensión cordial. Es precisamente en el estudio de esta dimensión cordial donde la neurociencia presenta un interesante vínculo con la ética, sin pretensión de fundamentarla. Los estudios sobre las condiciones naturales de posibilidad de lo moral, su funcionamiento, los factores que intervienen en el juicio moral, etc., revelan la indisociable presencia de la razón y de las emociones en el ejercicio de la moralidad. En este sentido, la ciencia, y más específicamente la neurociencia, tiene mucho que decir sobre las condiciones en las que se produce todo diálogo y toda argumentación posible y válida. Si bien la información ofrecida por estos estudios científicos no permite ofrecer criterios justificativos, sí que permite comprender la situación en la que el diálogo que busca esos criterios se produce. Cortina advierte de que las aportaciones de la ciencia en relación a los vínculos entre las costumbres y las emociones no deben servir de motivo para apelar a estas últimas de forma dogmática como fundamento de la ética, tal es la pretensión de algunos autores que, apelando a la dimensión emocional de nuestra naturaleza, intentan fundamentar científicamente la ética, de forma falaz (Spencer 1964, 1883; Richards, 1986; Arnhart 1998; Singer, 1981; Casebeer 2003, 34; Flanagan; Cartwright 2000, 328; McShea y McShea 1990, 1999; Churchland 2013; Curry 2006, 235, 237-238; Mora, 2007; Gazzaniga, 2006; Greene, 2013).

La aportación de Cortina, en este punto, resulta de gran relevancia para la neuroética, siendo clave a la hora de proponer una metodología adecuada para una neuroética crítica, siendo este uno de los objetivos de nuestra investigación. Esta aportación consiste en la idea de que la filosofía moral precisa una comprensión profunda de la situación natural de los agentes morales, solo así puede comprenderse la situación en la que se produce el diálogo normativo, que es al mismo tiempo racional y emotivo-estimativo, en virtud de la naturaleza de quienes participan en el diálogo.

Muestra de ello es el hecho de que todos los elementos cordiales que Cortina señala como implicados en el diálogo han sido estudiados desde las ciencias empíricas, en tanto que estructuras naturales de la moralidad o guardando alguna relación estrecha con tales estructuras. Tales elementos son: la obligación moral, que se produce como una ligatio motivada emocionalmente (Piaget, 1987; Kohlberg, 1992; Gilligan, 2003); el reconocimiento, que no se produce solo como condición pragmático-trascendental (Apel, 1985: 404; Habermas, 2006: 24-25), sino que se produce también en términos cordiales

o compasivos (Colmenares, 2005; Morgado, 2010; Evers, 2010; Eisenberg, 2000; Tagney, 2007; Damasio, 2008; Waal, 1997, 2007; Brosnan y Waal, 2003; Tomasello y Vaish, 2012; Churchland, 2013; Salles, 2015); el interés, que biológicamente está predispuesto de forma egocéntrica y endogámica (Waal, 1997, 2007; Brosnan y Waal, 2003; Evers, 2010; Churchland, 2013; Cortina, 2012a, 2017), pero puede configurarse de forma altruista o universal (Ayala y Cela Conde, 2001; Evers, 2010), ante la apreciación de lo valioso, la indignación ante lo injusto, y ante la estima por la propia vida y la de los otros (Kant, 1992; Korsgaard, 2000); las emociones, entre las que encontramos el sentimiento de simpatía, la compasión, el sentimiento de respeto, la reputación, así como facultades emocionales, tales como la empatía (Darwin, 1979; Waal, 1997, 2007; Eisenberg, 2000; Tagney, 2007; Moll, 2002; Greene; Rizzolatti, 2006; Morgado, 2010; Damasio, 2005; Damasio, 2008; Haidt, 2001; Churchland, 2013; Salles, 2015); la responsabilidad, que se nos presenta ya que podemos educar nuestras emociones, dada la plasticidad del cerebro (Sherman, 1999; Goleman, 1996; Evers, 2010; Cortina, 2009, 2017; Jácomo y Serrão, 2013); la capacidad estimativa, que presenta una base biológica pero puede darse en un sentido más amplio (Evers, 2010; Churchland, 2013; Damasio; Aranguren, 1994); la razón, que siempre se produce de forma vinculada a las emociones, ya sea en mayor o menor medida (Damasio, 2005, 2008; Evers, 2010; Greene; Haidt, 2001, 2013; Cortina); y otros elementos, tales como la conciencia moral, el sentimiento moral, la autoestima, el amor al prójimo, etc., que pueden considerarse como condiciones subjetivas de la receptividad en clave cordial (Cortina, 2009).

Directrices metodológicas para una neuroética de la razón cordial.

Los aspectos contemplados en el presente capítulo guardan una estrecha conexión, como hemos visto, con las diferentes cuestiones que hemos analizado a lo largo de la investigación. Atendiendo a estas es posible confeccionar, bajo una perspectiva cordial, una propuesta de método adecuada para la neuroética. Esta propuesta metodológica integra la ciencia y la ética bajo la perspectiva dialógica, presentándose cada una de ellas como un discurso con su propio método y competencias, siendo ambos producidos en contextos comunicativos. Contempla, también, algunos aspectos metodológicos relevantes dada la comprensión que se ofrece del discurso científico y normativo. Además, aporta algunas recomendaciones de método derivadas del estudio y crítica de diferentes investigaciones relevantes para el campo de la neuroética. Todo ello atendiendo

a la relevancia de la ética de la razón cordial en el ámbito de las neurociencias. En este apartado, retomando y compilando algunas de las cuestiones metodológicas que hemos tratado a lo largo de la investigación, confeccionaremos la propuesta de un método adecuado para una neuroética crítica en perspectiva cordial, al que hemos llamado *neuroética de la razón cordial*.

El método de una neuroética de la razón cordial, en la medida en que se encuadra bajo la perspectiva filosófica de las éticas dialógicas hará uso del método trascendental. Este permite reconocer que una dualidad de perspectivas metodológicas es posible, correspondiendo estas a la ciencia y a la ética, respectivamente. Ambas comparten sus condiciones trascendentales de posibilidad, en la medida en que ambas se producen como comunidades de comunicación. Este *a priori* de toda argumentación posible y válida que ambas comparten presenta una serie de presupuestos normativos *a priori*, entre los cuales encontramos la justificación racional de toda afirmación, el reconocimiento mutuo entre interlocutores, la disposición al diálogo, la suposición de mutua racionalidad entre interlocutores, el reconocimiento de una validez incondicionada de los enunciados resultantes del consenso en actos comunicativos, la libertad, la ausencia de coacción y de engaños, la imparcialidad, la publicidad, la igualdad de oportunidades, la actitud crítica hacia los demás y hacia uno mismo, la pretensión de universalidad, la apertura hacia una fusión de horizontes, la esperanza y el compromiso.

A estos presupuestos hay que sumar la razón, entendida como razón cordial. Es necesario atender al concepto de razón cordial, ya que el ejercicio de la razón se da siempre de forma encarnada, por lo cual las condiciones trascendentales de toda discusión han de incluir los aspectos cordiales de la razón y del diálogo, sin los cuales este último no puede producirse. Desde la propuesta cordial, las condiciones de posibilidad de toda argumentación posible y válida toman la forma de ciertos derechos y libertades, entre los cuales están el derecho de participación, el derecho a cuestionar cualquier afirmación, el derecho a expresar la posición, las necesidades y los intereses propios, el derecho a no ser obligado y a no decidir por coacción, el derecho a la vida, la libertad de conciencia, la libertad de expresión y de información, la libertad de religión y de opinión, la libertad de asociación, el derecho a participar en la deliberación sobre normas que me afectan, y los derechos económicos, sociales y culturales que garanticen unas condiciones materiales y culturales que permitan el diálogo en los términos descritos.

Todo discurso, sea el teórico, propio de la ciencia, o el normativo, propio de la ética, ha de producirse tomando la situación ideal de habla como idea regulativa. El consenso dialógico se producirá intersubjetivamente, como resultado los enunciados propuestos ante los que se asienta tendrán vigencia y validez, siempre revisable. La ciencia y la ética presentan, de este modo, un carácter crítico, que permitirá evitar el dogmatismo y la ideología, favoreciendo la emancipación.

Es preciso indicar, además, los aspectos específicos de cada uno de estos discursos. El discurso teórico, correspondiente a la ciencia, presenta un paradigma naturalista, siendo su competencia atender el ámbito descriptivo y explicativo en relación al mundo natural. Sus enunciados presentan pretensión de verdad respecto a ese mundo, la vigencia de las teorías científicas es alcanzada mediante la investigación a través del método científico, cuyos resultados, en última instancia, han de contar con el asentimiento de la comunidad científica, entendida como comunidad de argumentación. En tanto que comunidad de comunicación, la ciencia no hace uso solamente de la perspectiva epistémica de tercera persona, sino también de la del partícipe, en la que la primera y la segunda persona se hayan implicadas. A los presupuestos normativos de toda argumentación hemos de sumar uno más en el caso de la ciencia, el supuesto de que la búsqueda de la verdad es deseable. En lo que concierne a la ciencia en relación con el estudio de lo moral, su objeto serán los aspectos naturales de la moralidad, que corresponden a las bases estructurales de esta.

En cuanto al discurso normativo, propio de la filosofía moral o ética, este presenta un paradigma filosófico, siendo el método trascendental el que consideramos como más adecuado para desempeñar las funciones normativas de la ética. Entre las competencias de esta encontramos el ámbito justificativo de la moralidad. Sus enunciados tienen pretensión de validez normativa, para que estos adquieran vigencia y validez precisan que haya un asentimiento intersubjetivo en el consenso derivado de la comunidad de comunicación compuesta por los afectados por la norma. La perspectiva epistémica que toma parte en este proceso es la del partícipe, compuesta por la primera y la segunda persona. Su objeto de estudio son los aspectos normativos de la moralidad, debiendo atender al problema de fundamentación, de modo que se ofrezcan criterios justificativos en relación a aquellos contenidos morales que resistan la crítica.

Una mala distinción de estos dos tipos de discurso, así como de sus competencias y objetos de estudio pueden conducir a un reduccionismo naturalista. Para evitarlo hay que ser crítico con las posturas reduccionistas que abogan por la posibilidad de una ética naturalizada o de una ciencia éticamente normativa. Para ello es preciso conocer y evitar los errores argumentativos que conducen a estas posturas (principalmente: falacia de Hume o del salto es-debe; falacia de Moore o definicionista [que puede darse de dos modos: falacia naturalista y falacia metafísica]; falacia de ambigüedad por equívoco o por homonimia; y falacia por transposición categorial). También resulta necesario distinguir entre bases (estructurales) y fundamentos (éticos) de la moralidad, así como entre estructuras (innatas, aunque maleables) y contenidos (contingentes) de lo moral. Por último, es preciso reconocer en qué es relevante la naturaleza humana para lo moral. Puede resultar relevante como condición de posibilidad de la agencia moral, como una restricción de viabilidad sobre la moralidad, como restricción de lo que es bueno para nosotros, como una compleja red de elementos interdependientes (esta distinción es clave en el debate sobre la mejora humana) y como motivo para atribuir un valor moral especial a cierto tipo de seres. Pero no es relevante como fundamento de normas morales.

Podemos señalar, además, algunos aspectos metodológicos específicos en el campo de la neuroética. En primer lugar atenderemos a aquellos que conciernen a su vertiente naturalista, cuyo objeto de estudio, como ya hemos indicado, son los aspectos naturales de la moralidad. Es necesario incidir en la necesidad de atender a las emociones, ya que históricamente no se ha hecho de modo adecuado. También es preciso reconocer la conciencia y la cognición como posibles objetos de investigación científica. En relación a su estudio la perspectiva de primera persona y los informes personales son relevantes, sobre todo en lo que atañe a la conciencia, ya que se experimenta siempre en primera persona. Se ha de reconocer, además, un monismo ontológico, estando lo mental y lo físico insoslayablemente entrelazado. Por último, se ha de evitar el eliminativismo (que puede ser conceptual, contextual y ontológico), la psicofobia y el cognitivismo ingenuo. En segundo lugar, atenderemos a aquellos aspectos metodológicos que conciernen al ámbito justificativo de lo moral. En este sentido es necesario señalar el papel de la filosofía moral en el ámbito de la neuroética, reconociendo que tiene vigencia y es necesaria, la ciencia no puede desplazarla ya que, tal y como hemos mostrado a lo largo de la investigación, el reduccionismo naturalista no resiste la crítica. La ciencia, a lo sumo, puede informar a la ética ofreciendo herramientas para perseguir los fines que esta

contempla. Entre las aportaciones que la ciencia ofrece a la ética cabe destacar que aquella facilita una comprensión de lo moral en su nivel estructural, favorece una educación y una neuroeducación moral más eficiente, permite comprender los elementos individuales y sociales en los que se produce el diálogo normativo y revela el carácter cordial de la razón.

En relación a este carácter cordial de la razón, hemos de indicar aquellos aspectos metodológicos que han de ser tenidos en cuenta en la neuroética. Esta confirma y respalda la tesis de que existe un vínculo insoslayable entre la razón y los aspectos cordiales, como las emociones o las capacidades estimativas. La neurociencia, además, ofrece a la ética información sobre las condiciones naturales en las que se produce el diálogo normativo, confirmando el nexo entre razón, por un lado, y emoción y capacidades estimativas, por otro. También permite una descripción y una explicación naturalista de los elementos cordiales implicados en el diálogo normativo (obligación moral, reconocimiento, emociones, responsabilidad, capacidades estimativas, conciencia moral, sentimiento moral, autoestima, amor al prójimo, etc.), aunque no por ello hemos de reducir la comprensión de estos elementos a las condiciones naturales que son su base. Por último, la neurociencia también ofrece información sobre el modo en que las estructuras naturales pueden modificarse para adecuarse mejor a las recomendaciones de la ética.

Finalmente, es conveniente hacer algunas recomendaciones metodológicas específicas para ciertos campos de investigación propios de la neuroética, ya que el análisis de las principales propuestas concernientes a ellos ha revelado su carácter problemático. Atenderemos aquí a las cuestiones del libre albedrío y del juicio moral.

Atendamos, en primer lugar, a la investigación naturalista del libre albedrío. En su estudio hemos de tener en cuenta que el experimento plantea una situación artificial, hay que ser cautos al extender los resultados a otros contextos. Existe riesgo, además, de adulterar el resultado, por ejemplo entrenando a los sujetos, que pueden mecanizar algunos aspectos de la decisión. También es preciso aclarar la terminología (ej. la decisión libre, ¿es necesariamente consciente?) y los criterios de cuantificación (¿puede cuantificarse la decisión en instantes?). Se han de especificar, además, los supuestos de la investigación (ej. la existencia del yo suele estar presupuesta, pero no demostrada). También se ha de descartar el epifenomenalismo dado que las razones influyen en la acción. Además de los problemas asociados al estudio del libre albedrío, hemos de

preguntarnos si este estudio es posible para la ciencia, en este sentido se ha de considerar la posibilidad de un agnosticismo en relación a esta cuestión. La existencia o inexistencia del libre albedrío, entendido como hecho natural, quizás sea una realidad incognoscible para las ciencias empíricas, por no poder someterse a medición a través de su método. Finalmente, se recomienda hablar de condiciones, y no de causas. La dificultad que supone tratar el tema de la libertad desde el paradigma de causalidad propio de las ciencias empíricas puede evitarse sustituyendo el concepto de causa por el de condición, hablando así en términos de probabilidad en lugar de hablar en términos de necesidad, lo cual permite, además, contemplar los factores que imponen límites a la acción.

Atendamos, en segundo lugar, a la investigación del juicio moral. El juicio moral es un objeto de estudio adecuado para la investigación neurocientífica. En su estudio es preciso tener en cuenta la complejidad de los procesos implicados en el juicio moral, evitando modelos demasiado simplistas, como el de Greene. En cuanto a las condiciones del experimento, estas pueden condicionar los resultados del mismo. El experimento, además, ofrece información limitada al contexto de este, extrapolar conclusiones ha de hacerse con cautela. En cuanto a las técnicas de neuroimagen, la interpretación de la información que estas ofrecen ha de realizarse de forma apropiada. El uso de dilemas también puede resultar problemático, las situaciones que plantean son poco realistas, lo cual dificulta su utilidad para la investigación y para la aplicación, pudiendo adulterar los resultados en la investigación. Es necesario, además, hacer uso de un rigor terminológico que permita evitar anfibologías y errores lógicos. Asimismo, es preciso evitar errores en las premisas que puedan comprometer la verdad de las conclusiones. Por último, se ha de evitar el reduccionismo naturalista, que Greene comete por varias vías (falacia por homonimia y transposición categorial, tomar las intuiciones como éticamente normativas y tomar la psicología moral como aval de la filosofía moral).

Todas estas recomendaciones reúnen los aspectos metodológicamente más relevantes que hemos estudiado a lo largo de nuestra investigación, permitiendo sintetizar en una sola propuesta un método adecuado para una neuroética crítica en perspectiva cordial, a la que hemos acuñado como neuroética de la razón cordial (ver Tabla 5).



Tabla 5. Aspectos generales de una neuroética de la razón cordial como propuesta metodológica para una neuroética crítica en clave cordial.

CONCLUSIONES

El nacimiento de la neuroética como una rama del conocimiento interdisciplinar planteó una serie de retos y problemas que han de ser atendidos y resueltos. Entre estos encontramos la necesidad de establecer un método adecuado, que permita que las ciencias empíricas y la filosofía moral puedan operar conjuntamente, sin que ninguna de ellas colonice las competencias de la otra. De este modo la neuroética podrá configurarse de forma verdaderamente interdisciplinar, en la que neurociencias y ética ocupen el papel que les corresponde, de tal modo que la disciplina no sea tan solo una investigación neurocientífica sobre lo moral, evitando, además, el riesgo de intentar fundamentar la ética desde las ciencias empíricas, tal y como algunos autores han pretendido.

Partiendo de la hipótesis de que las cuestiones fundamentales de la ética no pueden resolverse desde el paradigma naturalista de las ciencias empíricas, la presente investigación planteaba como su objetivo principal articular una propuesta metodológica verdaderamente interdisciplinar destinada a facilitar un adecuado ejercicio de la neuroética, en la que la filosofía y la ciencia puedan compaginarse. De este modo nuestra propuesta de método podría permitir una neuroética crítica, en la que la filosofía moral sería reconocida en sus competencias, todo ello desde una perspectiva dialógica y cordial.

A lo largo de la investigación hemos analizado desde una perspectiva crítica las aportaciones más relevantes para el campo de la neuroética. La revisión crítica de las propuestas científicas, tanto en neurociencias como en otras disciplinas relevantes para el estudio de lo moral, tales como la biología, la psicología, la sociología y la etología, entre otras, han permitido, por un lado, la comprensión de los aspectos que posibilitan la moralidad estructuralmente y, por otro, han posibilitado detectar los aspectos metodológicamente adecuados así como los metodológicamente problemáticos, en relación a los cuales hemos tenido la ocasión de reflexionar. La intervención de la filosofía, en este sentido, se presenta como necesaria, y permite establecer un marco crítico de reflexión y análisis que facilita someter a revisión no solo los contenidos que la ciencia ofrece, sino el modo en que lo hace. Tras este análisis filosófico del quehacer científico en el estudio de la moralidad, dos son las consideraciones más importantes que hemos de hacer. La primera de ellas es que la ciencia precisa ser revisada desde un análisis

crítico que la filosofía puede ofrecer, de tal modo que su método y sus competencias puedan ser justificados adecuadamente, cribando los criterios metodológicos que guían su labor. La segunda de ellas es que la ciencia está inhabilitada para abordar todas las dimensiones de la moralidad, ya que hay cuestiones que escapan a sus competencias, en la medida en que estas pertenecen al ámbito de la filosofía moral.

Se hace necesario, por tanto, ofrecer un marco filosófico que permita una comprensión conjunta de la ciencia y la filosofía moral, para así articular la posibilidad de una neuroética verdaderamente interdisciplinar. En este punto hemos atendido a las propuestas pertenecientes al procedimentalismo dialógico. Esta tradición de pensamiento ofrece algo más que una teoría ética, ofrece una comprensión de las condiciones de posibilidad de toda argumentación posible, englobando bajo esta categoría tanto el discurso teórico de la ciencia como el discurso normativo de la ética. El marco filosófico ofrecido por esta escuela de pensamiento es adecuado para la propuesta metodológica que la investigación persigue, pero es insuficiente desde las propuestas dialógicas más clásicas, en la medida en que estas son excesivamente racionalistas. Incorporar el componente cordial de la razón se presenta como una necesidad para la neuroética en dos sentidos. En primer lugar, los estudios neurocientíficos revelan el insoslayable vínculo entre la razón y lo cordial, compuesto principalmente por las emociones y las capacidades estimativas. En segundo lugar, la crítica filosófica enfocada desde una *ethica cordis*, haciendo uso del método trascendental, revela el carácter cordial de la razón como una de las condiciones de posibilidad de todo discurso argumentativo, incluyendo el teórico y el normativo. A todo ello hemos sumado la crítica de dos propuestas metodológicas en el ámbito de la neuroética, la neurofilosofía práctica, con un carácter marcadamente naturalista y metodológicamente problemático, y el materialismo ilustrado, del cual hemos tomado algunas aportaciones interesantes para nuestra propuesta, por ser un modelo metodológico que permite compatibilizar la labor científica y la filosófica.

En la articulación de nuestra propuesta de método, que hemos llamado *neuroética de la razón cordial*, hemos perseguido, además, cinco objetivos específicos. El primero de ellos aborda el problema de la naturalización de la moral en las neurociencias, en relación a esta cuestión hemos tratado de aclarar si una naturalización de este tipo es posible, y de serlo, cómo. Tras el análisis de las diferentes investigaciones científicas y neurocientíficas, podemos concluir que una naturalización de la moral desde las

neurociencias es posible parcialmente. La moralidad acontece como un hecho natural, pero todo lo concerniente a lo moral no es reductible a ese hecho. Las condiciones naturales de la moralidad permiten abrir el espacio de la ética, entendida como una reflexión filosófica que versa sobre los criterios que justifican la acción, las normas, los valores, las ideas y las virtudes morales. En este sentido, el paradigma naturalista puede abordar el estudio de las condiciones materiales y estructurales que son la base de la moralidad, así como su funcionamiento, todo ello supone un objeto de estudio apropiado para ciencia y para las neurociencias. Sin embargo, en lo concerniente a la dimensión justificativa de la moralidad, el paradigma naturalista es insuficiente ya que su cometido es explicar y describir el mundo, corresponde a la ética ocuparse de los criterios de justificación y fundamentación de lo moral, los cuales escapan al método científico por no ser reductibles a lo natural.

El segundo de los objetivos específicos que nuestra investigación ha perseguido consiste en especificar los aspectos generales de la ciencia y de la ética, atendiendo a la relación que guardan, a su método, a las competencias que les corresponden y a sus objetos de estudio, siendo todo ello de interés a la hora de ofrecer una propuesta de método para la neuroética. Como ya hemos expuesto, las diferentes propuestas de ética dialógica a las que hemos atendido durante la investigación ofrecen no solo una teoría ética, sino los recursos conceptuales necesarios para establecer un marco que englobe tanto al discurso teórico como al normativo, en la medida en que ambos se configuran en el seno de comunidades de argumentación. Al teórico, propio de las ciencias, corresponde la tarea de explicar y describir el mundo natural, a través del método científico. Al normativo, propio de filosofía moral, corresponde la tarea de fundamentación y justificación de normas, valores, principios, virtudes e ideas morales, siendo el método trascendental el que nuestra propuesta suscribe, entre todos aquellos que la filosofía ofrece. Cabe señalar que este método engloba al discurso científico en tanto que se produce en el seno de una comunidad de comunicación.

El tercero de los objetivos específicos planteados en nuestra investigación propone analizar las aportaciones más relevantes en el campo de la neuroética, incluyendo no solo las aportaciones científicas, sino también las de carácter más filosófico, para determinar qué aspectos resultan metodológicamente adecuados y cuáles metodológicamente problemáticos en el ejercicio de esta disciplina. A lo largo de nuestro estudio hemos

tenido ocasión de atender a diversas propuestas y autores, reflexionando críticamente en torno a las aportaciones que ofrecen. Muchas de estas propuestas ofrecen criterios de investigación interesantes y adecuados para el ejercicio conjunto de la ciencia y de la filosofía moral, sin embargo otras han resultado ser muy problemáticas. Los problemas metodológicos que hemos podido encontrar conciernen a cuestiones como el diseño de los experimentos, la interpretación de los datos, la terminología utilizada, los errores lógicos cometidos, el reduccionismo naturalista, los objetos de estudio que pueden abordarse desde los distintos discursos, etc. Tras someter a revisión las aportaciones que han sido objeto de nuestro estudio, hemos podido extraer una serie de recomendaciones metodológicas que han de ser tenidas en cuenta para poner en ejercicio la neuroética de forma adecuada. Nuestra propuesta de una neuroética de la razón cordial recoge los resultados cosechados en la consecución de este objetivo.

El cuarto de los objetivos específicos planteados en nuestra investigación consiste en ofrecer una descripción y una explicación científica y neurocientífica de los aspectos naturales de la moralidad. Este objetivo se centra, por tanto, en la dimensión estructural de lo moral, y concierne únicamente al ámbito descriptivo y explicativo de la ciencia. Para alcanzar este objetivo hemos atendido, en primer lugar, a diversas propuestas científicas acerca del origen filogenético de la moralidad, su desarrollo ontogenético y su funcionamiento cerebral. Partiendo de todo ello hemos ofrecido una hipótesis acerca de las condiciones naturales de posibilidad de la agencia moral, entendida como la facultad psicobiológica que permite la moralidad. Nuestra hipótesis recoge las aportaciones de diversos autores, pero aporta un elemento novedoso. Atendiendo a diversas contribuciones científicas, planteamos la aparición de la moralidad en el ser humano como un subproducto evolutivo, derivado de la concurrencia de determinadas facultades y elementos cognitivos y emocionales, habiendo alcanzado estos determinado nivel de complejidad. Como elemento novedoso aportamos una propuesta específica acerca de cómo se produce esta concurrencia de elementos y cómo esta conduce a la aparición de lo moral. Tras proponer esta tesis, a lo largo de la investigación se amplía el estudio de la dimensión natural de la moralidad, atendiendo, entre otras cuestiones, al estudio del juicio moral y de su funcionamiento, a la distinción entre estructuras y contenidos de lo moral, al desarrollo de las estructuras morales durante la epigénesis, etc. Podemos comprobar que nuestra hipótesis inicial sobre las condiciones naturales de la moralidad es coherente con las diferentes aportaciones que a lo largo de la investigación van ampliando la

exposición de los aspectos naturales de la moralidad, completando estas últimas la consecución de este objetivo.

Por último, el quinto objetivo específico consiste en ofrecer criterios para una fundamentación ética de contenidos morales desde la perspectiva de las éticas dialógicas. Si bien, como ya hemos indicado, esta corriente de pensamiento permite situar tanto al discurso teórico como al normativo como resultado de comunidades de comunicación, este objetivo atañe tan solo al segundo de estos discursos. Del mismo modo que el objetivo anterior se ocupaba únicamente de los aspectos naturales de la moralidad, este objetivo trata de dar cuenta únicamente de los aspectos normativos de esta. A lo largo de la investigación hemos atendido a los autores más relevantes de esta corriente, poniendo de relieve que las propuestas más clásicas de la misma son deficitarias en la medida en que resultan excesivamente racionalistas. La incorporación de los aspectos cordiales al concepto de razón permite una comprensión más profunda de las condiciones en las que se produce el diálogo. Por ello, suscribimos los aspectos generales de las propuestas dialógicas clásicas, según las cuales una norma adquiere validez y vigencia mediante el consenso intersubjetivo producido en la comunidad de afectados por la norma. Pero reconocemos la necesidad de actualizar tales propuestas desde el concepto de razón cordial. Para comprender adecuadamente las condiciones de validez y posibilidad de toda argumentación, incluyendo el discurso normativo, no se puede ignorar que la razón es siempre ejercida como razón cordial.

A la luz de estas consideraciones, podemos concluir que los objetivos de la investigación han sido alcanzados de forma satisfactoria a lo largo de nuestro estudio. A través de este no solo hemos ofrecido una propuesta de método, sino que la hemos puesto en ejercicio, mostrando que es posible una neuroética crítica verdaderamente interdisciplinar.

CONCLUSIONS (english version)

The birth of neuroethics as a field of interdisciplinary knowledge posed a series of challenges and problems that must be addressed and resolved. Among them we find the need to establish an adequate method, which allows the empirical sciences and moral philosophy to operate jointly, without any of them colonising the competences of the other. In this way, neuroethics can be configured in a truly interdisciplinary way, in which neurosciences and ethics occupy the role that corresponds to them, in such a way that the discipline is not only a neuroscientific research on the morality, also avoiding the risk of trying to base ethics from the empirical sciences, as some authors have claimed.

Starting from the hypothesis that the fundamental questions of ethics can not be solved from the naturalist paradigm of the empirical sciences, this research proposed as its main objective to articulate a truly interdisciplinary methodological proposal destined to facilitate an adequate practice of neuroethics, in which philosophy and science can come together. In this way, our methodological proposal could allow a critical neuroethics, in which moral philosophy would be recognised in its competences, all from a dialogical and cordial perspective.

Throughout the research, we have analysed the most relevant contributions to the neuroethics field from a critical perspective. The critical review of the scientific proposals both in neurosciences and in other disciplines relevant to the study of morality, such as biology, psychology, sociology and ethology among others, have allowed, on one hand, the understanding of the aspects that enable morality structurally and, on the other, have made possible to detect the methodologically appropriate aspects as well as the methodologically problematic ones, in relation to which we have had the opportunity to reflect on. The intervention of philosophy, in this sense, is presented as necessary and allows to establish a critical framework of reflection and analysis that facilitates reviewing not only the contents that science offers, but the way in which it does so. After this philosophical analysis of the scientific work in the study of morality, two are the most important considerations we have to make. The first one is that science needs to be revised from a critical analysis which philosophy can offer, in such a way that its method and its competences can be adequately justified, by screening the methodological criteria that guide its work. The second one is that science is incapable of addressing all the

dimensions of morality, since there are issues that escape its competences, insofar as they belong to the field of moral philosophy.

Therefore, it is necessary to offer a philosophical framework that allows a joint understanding of science and moral philosophy in order to articulate the possibility of a truly interdisciplinary neuroethics. At this point, we have heeded the proposals pertaining to the dialogical procedure. This tradition of thought offers something more than an ethical theory. It offers an understanding of the conditions of possibility of all possible argumentation, encompassing both the theoretical discourse of science and the normative discourse of ethics under this category. The philosophical framework offered by this school of thought is adequate for the methodological proposal pursued by the research, but it is insufficient from the more classical dialogical proposals, to the extent that they are excessively rationalist. Incorporating the cordial component of reason is presented as a need for neuroethics in two senses. In the first place, neuroscientific studies reveal the unavoidable link between reason and cordiality, composed mainly of the emotions and the estimative capacities. In the second place, philosophical criticism focused from an *ethica cordis*, making use of the transcendental method, reveals the cordial nature of reason as one of the conditions of possibility of all argumentative discourse, including the theoretical and the normative ones. To all this, we have added the criticism of two methodological proposals in the field of neuroethics, practical neurophilosophy, with a markedly naturalistic and methodologically problematic character, and informed materialism, from which we have taken some interesting contributions to our proposal, since it is a methodological model that allows the scientific and philosophical work to be compatible.

In the articulation of our methodological proposal which we have called neuroethics of cordial reason, we have also pursued five specific objectives. The first one addresses the problem of the naturalization of morality in the neurosciences. In relation to this issue we have tried to clarify whether a naturalization of this type is possible, and if so, how. After analysing the different scientific and neuroscientific investigations, we can conclude that a naturalization of morality from the neurosciences is partially possible. Morality happens as a natural fact, but everything concerning morality is not reducible to that fact. The natural conditions of morality allow opening the space of ethics, understood as a philosophical reflection that deals with the criteria that justify moral actions, norms, values, ideas and virtues. In this sense, the naturalist paradigm can address the study of

the material and structural conditions that are the basis of morality, as well as its functioning, all this is an object of study appropriate for science and for neurosciences. However, regarding the justificatory dimension of morality, the naturalist paradigm is insufficient since its mission is to explain and describe the world, it is up to ethics to deal with the criteria of justification and foundation of morality, which escape to the scientific method for not being reducible to the naturalistic paradigm.

The second specific objective that our research has pursued is to specify the general aspects of science and ethics according to their relationship, their method, the competences that correspond to them and their objects of study; being all this of interest when offering a methodological proposal for neuroethics. As we have previously stated, the different dialogic ethics proposals that we have studied during the research offer not only an ethical theory, but also the conceptual resources necessary to establish a framework that encompasses both the theoretical and the normative discourse, to the extent that both are configured within the communities of argumentation. The task of explaining and describing the natural world through falls to the theoretical discourse, typical of the sciences by using the scientific method. The task of the foundation and justification of the moral norms, values, principles, virtues and ideas is responsibility of the normative discourse, typical of moral philosophy, being the transcendental, among all the others methods that philosophy offers, the one that our proposal subscribes. It should be noted that this method encompasses scientific discourse insofar as it occurs within a community of communication.

The third specific objective brought up in our research is to analyse the most important contributions in the field of neuroethics, including not only the scientific contributions but also the more philosophical ones, in order to determinate which aspects are methodologically adequate and which ones are methodologically problematic in the practice of this discipline. Throughout our study, we have paid attention to several proposals and authors, critically reflecting on the contributions they offer. Many of these proposals offer interesting and adequate research criteria for the joint practice of science and moral philosophy, however, others turned out to be very problematic. The methodological problems that we found are in regards issues such as the design of the experiments, the interpretation of the data, the terminology used, the logical errors made,

the naturalist reductionism, the objects of study that can be addressed from each discourse, etc. After reviewing the contributions that have been the object of our study, we have extracted several methodological recommendations that must be taking into account in order to put neuroethics in practice correctly. Our proposal of a neuroethics of cordial reason collects the results gathered in the achievement of this objective.

The fourth specific objective posed in our research is to offer a description and a scientific and neuroscientific explanation of the natural aspects of morality. This objective focuses, therefore, in the structural dimension of morality and it only concerns the descriptive and explanatory field of science. To reach this objective we have heeded, first of all, various scientific proposals about the phylogenetic origin of morality, its ontogenetic development and its cerebral functioning. On this basis, we have offered a hypothesis about the natural conditions of possibility of moral agency, understood as the psychobiological faculty that allows morality. Our hypothesis collects the contributions of several authors, but provides a novel element. Paying attention to diverse scientific contributions, we propose the appearance of morality in the human being as an evolutionary byproduct, derived from the concurrence of certain faculties and some cognitive and emotional elements, having these reached a certain level of complexity. As a novel element we provide a specific proposal about how this concurrence of elements occurs and how this leads to the appearance of morality. After proposing this thesis, throughout the investigation, the study of the natural dimension of morality is extended, taking into account, among other issues, the study of moral judgment and its functioning, the distinction between structures and contents of morality, the development of moral structures during epigenesis, etc. We can prove that our initial hypothesis about the natural conditions of morality is coherent with the different contributions that throughout the investigation extend the exposure of the natural aspects of morality, completing the achievement of this objective.

Lastly, the fifth specific objective is to offer criteria for an ethical foundation of moral contents from the perspective of dialogical ethics. Although, as we have already indicated, this current of thought allows both the theoretical and the normative discourse to be put as a result of communication communities, this objective concerns only the normative one. Just as the previous objective deals only with the natural aspects of

morality, this objective tries to account only for the normative aspects of it. Throughout the investigation we have paid attention to the most relevant authors of this current, emphasising that its most classic proposals are deficient to the extent they are excessively rationalistic. The incorporation of cordial aspects into the concept of reason allows a deeper understanding of the conditions in which dialogue occurs. For this reason, we endorse the general aspects of the classical dialogical proposals, according to which a norm acquires validity through the intersubjective consensus produced in the community affected by the norm. But we recognise the need to update such proposals from the concept of cordial reason. To properly understand the conditions of validity and possibility of any argument, including the normative discourse, the fact that reason is always exercised as cordial reason can not be ignored.

In the light of these considerations, we can conclude that the objectives of the investigation have been achieved in a satisfactory way throughout our research. Through the research, we have not only offered a methodological proposal but we have also used it, showing that a truly interdisciplinary neuroethics is possible.

BIBLIOGRAFÍA

- Albert, H. (1978). *Ética y metaética: el dilema de la filosofía moral analítica*. Valencia: Revista Teorema.
- Anderson, S.W. et al. (1999). Impairment of social and moral behavior related to early damage in human prefrontal cortex. *Nat. Neurosci.* 2: 1032-1037.
- Apel, K. O. (1985). El *a priori* de la comunidad de comunicación y los fundamentos de la ética. En Apel, K. O., *La transformación de la filosofía* (Vol. II), 341-413. Madrid: Taurus.
- Aranguren, J.L.L. (1994). *Ética*. Madrid: Alianza Editorial.
- Aristóteles. (2002). *Ética a Nicómaco*. Madrid: Centro de Estudios Políticos y Constitucionales.
- Arnhart, L. (1998). *Darwinian Natural Right: The Biological Ethics of Human Nature*. Albany, NY: State University of New York Press.
- Ayala, F. J. (2006). Las raíces biológicas de la moral. En Ayala, F. J., *La evolución de un evolucionista*, 351-369. Valencia: Universidad de Valencia.
- Ayala, F. J. y Cela Conde, C. J. (2001). La filogénesis de la moral. En Ayala, F. J. y Cela Conde, C. J., *Senderos de la evolución humana*, 517-538. Madrid: Alianza Editorial.
- Banks, W. P. y Pockett, S. (2007). Benjamin Libet's work on the neuroscience of free will. En Velmans, M. y Schneider, S. (eds.), *Blackwell Companion to Consciousness*, 657-670. Blackwell Publishing.
- Baron, J. (1994). Nonconsequentialist decisions. *Behavioral and Brain Sciences*, 14: 1-42.
- Barrett, L., R. Dunbar y J. Lycett. (2002). *Human Evolutionary Psychology*. Princeton, NJ: Princeton University Press.

- Bentham, J. (2008). *Los principios de la moral y la legislación*. España: Editorial Claridad.
- Berker, A. S. (2009). The normative insignificance of neuroscience. *Philosophy & Public Affairs*, 37(4): 293-329.
- Bieri, P. (2002). *El oficio de ser libre*. Barcelona: Ariel.
- Binmore, K. (2005). *Natural justice*. Oxford: University Press.
- Blackburn, S. (1994). *Oxford Dictionary of Philosophy*. Oxford: Oxford University Press.
- Blackburn, S. (1998). *Ruling Passions*. New York: Oxford University Press.
- Blackburn, S. (2001). *Being Good: A Short Introduction to Ethics*. New York: Oxford University Press.
- Blair, R. J. (1995). A cognitive developmental approach to mortality: investigating the psychopath. *Cognition*, 57: 1-29.
- Blair, R. J. (2001). Neurocognitive models of aggression, the antisocial personality disorders, and psychopathy. *J. Neurol. Neurosurg. Psychiatry*, 71: 727-731.
- Blair, R. J. et al. (1997). The psychopathic individual: a lack of responsiveness to distress cues?. *Psychophysiology*, 34: 192-198.
- Bluhm, R. (2014). No Need for Alarm: A Critical Analysis of Greene's Dual-Process Theory of Moral Decision-Making. *Neuroethics*, 7: 299-316.
- Bock, W. J. (1980). The definition and recognition of biological adaptation. *American Zoologist*, 20: 217-227.
- Brandom, R. B. (1994). *Make it explicit*. Massachusetts: MIT Press.
- Brennan, G. y Pettit, P. (2004). *The economy of esteem: an essay on civil and political society*. Oxford: Oxford University Press.
- Brink, D. O. (1989). *Moral Realism and the Foundation of Ethics*. New York/Oxford: Oxford University Press.

- Brosnan, S. F. y Waal, F. B. (2003). Monkeys reject unequal pay. *Nature*, 425 (6955): 297-299.
- Buchanan, A. (2009). Human Nature and Enhancement. *Bioethics*, Vol. 23, N° 3: 141-150.
- Campbell, J. (1993). A Simple View of Colour. En J. Haldane y C. Wright (eds.), *Reality, Representation and Projection*. New York: Oxford University Press.
- Carnap, R. (1990). *Pseudoproblemas de la filosofía*. México D. F.: Universidad Autónoma de México.
- Cartwright, J. (2008). *Evolution and human behaviour*. London: Palgrave.
- Cartwright, J. (2010). Naturalising Ethics: The Implications of Darwinism for the Study of Moral Philosophy. *Sci & Educ*, 19: 407-443.
- Casebeer, W. D. (2003). *Natural Ethical Facts: Evolution, Connectionism, and Moral Cognition*. Cambridge, Massachusetts: MIT Press.
- Cayuela, A. (2012). The «evil ideology» description and critique of a policy against human nature (el «mal ideológico»: descripción y crítica de una política contraria a la naturaleza humana). *Cuadernos de Bioética XXIII*, 2012/1a: 71-82.
- Changeux, J. P y Connes, A. (1993). *Materia de reflexión*. Barcelona: 1993.
- Changeux, J. P. (1985). *El hombre neuronal*. Madrid: Espasa-Calpe.
- Changeux, J. P. (2010). *Sobre lo verdadero, lo bello y el bien: un enfoque neuronal*. Buenos Aires; Madrid: Katz.
- Changeux, J. P. y Ricoeur, A. (2001). *La naturaleza y la norma: lo que nos hace pensar*. México: Fondo de Cultura Económica.
- Chomsky, N. (1987). *Estructuras sintácticas*. México: Siglo XXI.
- Churchland, P. (2008). The Impact of Neuroscience on Philosophy. *Neuron*, 60: 409-411.

- Churchland, P. S. (2013). *El cerebro moral. Lo que la neurociencia nos dice sobre la moralidad*. Barcelona: Paidós.
- Colmenares, F. (2005). De laberintos sociales y de cómo salir de ellos: inteligencia social. En Guillén-Salazar, F. (ed.), *Existo, luego pienso: Los Primates y la Evolución de la Inteligencia Humana*, 75-128. Madrid: Ateles.
- Conill, J. (1991). *El enigma del animal fantástico*. Madrid: Tecnos.
- Conill, J. (2006). *Ética hermenéutica*. Madrid: Tecnos.
- Conill, J. (2012). La superación del naturalismo en Ortega y Gasset. *Isegoría*, 46: 167-192.
- Conill, J. (2015a). La intimidad corporal en la filosofía de Ortega y Gasset. *Isegoría*, 53: 491-513.
- Conill, J. (2015b). Naturaleza humana en perspectiva biohermenéutica. *Pensamiento*, vol.7, n. 269: 1249-1260.
- Conill, J. (2016). La intimidad corporal y sus bases neurobiológicas. *Pensamiento*, vol. 72, núm. 273: 789-807.
- Cortina, A. (2001). *Ética aplicada y democracia radical*. Madrid: Tecnos.
- Cortina, A. (2007). *Ética mínima: Introducción a la filosofía práctica*. Madrid: Tecnos.
- Cortina, A. (2009). *Ética de la razón cordial: educar en la ciudadanía en el siglo XXI*. Oviedo: Nobel, D. L.
- Cortina, A. (2010). *Ética sin moral*. Madrid: Tecnos.
- Cortina, A. (2012a). *Neuroética y neuropolítica: Sugerencias para la educación moral*. Madrid: Tecnos.
- Cortina, A. (2016). La conciencia moral desde una perspectiva neuroética. De Darwin a Kant. *Pensamiento*, vol. 72, núm. 273: 771-778.

- Cortina, A. (2017). *Aporofobia, el rechazo al pobre. Un desafío para la democracia*. Barcelona: Paidós.
- Cortina, A. (ed.) (2012b). *Guía Comares de Neurofilosofía Práctica*. Granada: Editorial Comares.
- Curry O. (2006). Who's afraid of the naturalistic fallacy?. *Evolutionary Psychology* 4: 234-247.
- Curry, O. (2005). *Morality as Natural History: An adaptationist account of ethics*. PhD Thesis. London: London School of Economics.
- Cushman, F. y Greene, J. D. (2012). Finding faults: How moral dilemmas illuminate cognitive structure. *Social Neuroscience*, vol. 7, nº 3: 269-279.
- Damasio, A. (2005). *En busca de Spinoza: Neurobiología de la emoción y los sentimientos*. Barcelona: Crítica.
- Damasio, A. (2008). *El error de Descartes: La emoción, la razón y el cerebro humano*. Barcelona: Crítica.
- Damasio, H. et al. (1994). The return of Phineas Gage: clues about the brain from the skull of a famous patient. *Science* 264: 1102-1105.
- Darwin, C. (1979). *El origen del hombre y la selección en relación al sexo*. Madrid: EDAF.
- Davidson, D. (2001). *Subjective, Intersubjective, Objective*. Oxford: Oxford University Press.
- Davidson, D. (2004). *Problems of Rationality*. Oxford: Oxford University Press.
- Dean, R. (2010). Does Neuroscience Undermine Deontological Theory?. *Neuroethics*, 3: 43-60.
- Delaney, T. (2009). Social spencerism. *Philosophy Now*, 71: 20-21.
- Dennett, D. (1995). *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. Simon & Schuster.

- Dewey, J. (1902). *The Evolutionary Method as Applied to Morality*. En J. A. Boydston, (ed.), *John Dewey: The Middle Works*, Vol. 2, 1899-1924. Carbondale and Edwardsville: Southern Illinois University Press.
- DiCarlo, C. (2000). Critical Notice of Anthony O'Hear's *Beyond Evolution: Human Nature and the Limits of Evolutionary Explanation*. *Biology and Philosophy*, 16: 117-130.
- Doris, J. M., y S. P. Stich. (2005). *As a Matter of Fact: Empirical perspectives on ethics*. En F. Jackson y M. Smith (eds.), *The Oxford Handbook of Contemporary Analytic Philosophy*. Oxford: Oxford University Press.
- Dummett, M. (1993). *The seas of language*. Oxford: Oxford University Press.
- Einsengberg, N. (2000). Emotion, Regulation and Moral Development. *Annual Review of Psychology*, 51: 665-697.
- Evers, K. (2010). *Neuroética. Cuando la materia se despierta*. Madrid: Katz Editores.
- Evers, K. (2015). Can We Be Epigenetically Proactive?. En T. Metzinger & J. M. Windt (eds). *Open MIND*: 13(T). Frankfurt am Main: MIND Group.
- Evers, K. (2017). Proactive epigenesis and ethics — Response. *EMBO Reports*, vol. 18: 1272-1272.
- Evers, K. y Changeux, J. (2016). Proactive epigenesis and ethical innovation. *EMBO Reports*, vol. 17: 1361-1364.
- Evers, K. (2016). The contribution of neuroethics to international brain research initiatives. *Nature Reviews Neuroscience*, vol. 18: 1-2.
- Evers, K., Salles, A. y Farisco, M. (2017). Theoretical framing of neuroethics: the need for a conceptual approach. *Debates about neuroethics*. E. Racine y J. Aspler (eds.). Springer.
- Farah, M. J. (2012). Neuroethics: The Ethical, Legal, and Societal Impact of Neuroscience. *Annual Review of Psychology*, 63: 571-91.

- Farber, P. L. (1994). *The temptations of evolutionary ethics*. Berkeley/Los Angeles/London: University of California Press.
- Farisco, M. y Evers, K. (2017). The ethical relevance of unconcious. *Philosophy Ethics and Humanities in Medicine*, vol. 12.
- Farisco, M., Evers, K. y Salles, A. (2016). Big Science, Brain Simulation and Neuroethics. *AJOB Neuroscience*, 7(1): 28-30.
- Farisco, M., Salles, A. y Evers, K. (2018). Neuroethics: A Conceptual Approach. *Cambridge Quarterly of Healthcare ethics*. (forthcoming October 2018)
- Fichte, J. G. (1910). *Werke*. Leipzig: Fritz Medicus.
- Flanagan O., H. Sarkissian y D. Wong. (2008). Naturalizing ethics. The evolution of morality: adaptations and innateness. En W. Sinnott-Armstrong W. (ed.), *Moral psychology*, Vol. 1, 1-25. Cambridge, Massachusetts: MIT Press.
- Flanagan, O. (1996). Ethics naturalized: ethics as human ecology. En M. F. Larry May y Clark Andy (eds.), *Mind and morals*, 19-43. Cambridge, Massachusetts: MIT Press.
- Flew, A. (1967). *Evolutionary Ethics*. London: Macmillan.
- Flew, A. (1978). From Is to Ought. En A. L. Caplan (ed.), *The Sociobiology Debate: Readings on ethical and scientific issues*, 142-162. New York: Harper and Row.
- Forbes, C. E. y Grafman, J. (2010). The Role of the Human Prefrontal Cortex in Social Cognition and Moral Judgment. *Annual Review of Neuroscience*, 33: 299-324.
- Frankena, W. K. (1939). The Naturalistic Fallacy. *Mind*, 48: 464-477.
- Frankena, W. K. (1957). Ethical Naturalism Renovated. *The Review of Metaphysics*, 10: 457-473.
- Frankena, W. K. (1973). *Ethics*. Prentice Hall, Englewood Cliffs.
- Frege, G. (1892). *Zeitschrift für Philosophie und philosophische Kritik*, Nueva Serie, N° 100: 25-50.

- Frith C.D. y Frith, U. (2008). Implicit and Explicit Processes in Social Cognition. *Neuron*, 60: 503-510.
- Fruteau, C., Voelkl, B., van Damme, E. y Noe, R. (2009). Supply and demand determine the market value of food providers in wild vervet monkeys. *PNAS*, Vol. 106, N° 29: 12007-12012.
- Gadamer, H. G. (1991). *Verdad y método: fundamentos de una hermenéutica filosófica*. Salamanca: Sígueme.
- García Damboreana, R. (2000). *Uso de razón: diccionario de falacias*. Madrid: Bbiblioteca Nueva.
- García-Marzá, D. (1992). *Ética de la justicia. J. Habermas y la ética del discurso*. Madrid: Tecnos.
- García-Marzá, D. (2005). Sentimientos virtuosos. El papel de los sentimientos en la vida moral. En *Diálogo filosófico*, N° 62: 241-256.
- García-Marzá, D. (2012). Neuropolítica: una mirada crítica sobre el neuropoder. En Cortina A. (ed.), *Neurofilosofía práctica*, 77-97. Granada: Comares.
- García-Marzá, D. (2013). Neuropolítica y democracia: un diálogo necesario. En *Daimon. Revista Internacional de Filosofía*, N° 59: 171-178.
- García-Marzá, D. (2016). Neuroética aplicada: las consecuencias prácticas del neuropositivismo. En *Pensamiento*, Vol. 72, N° 273: 881-900.
- Gaulin, S. J. C. y D. H. McBurney. (2001). *Psychology: An Evolutionary Approach*. NJ: Prentice Hall, Upper Saddle River.
- Gazzaniga, M. S. (2006). *El cerebro ético*. Barcelona: Paidós.
- Gewirth, A. (1993). How Ethical Is Evolutionary Ethics?. En D. Nitecki y M. Nitecki (eds.), *Evolutionary Ethics*, 241-256. Albany: SUNY Press.
- Gibbard, A. 1990. *Wise Choices, Apt Feelings: A Theory of Normative Judgement*. Cambridge, Massachusetts: Harvard University Press.

- Gilligan, C. (2003). *In a different voice: psychological theory and women's development*. Cambridge (MA): Harvard University Press.
- Goldberg, E. (2002). *El cerebro ejecutivo: lóbulos frontales y mente civilizada*. Barcelona: Crítica.
- Goleman, D. (1996). *Inteligencia emocional*. Barcelona: Kairos.
- Gollwitzer, P. (1999). Implementation intentions. *American psychologist*, N° 54: 493-503.
- González Esteban, E. (2004). La filosofía como brújula del presente. *Recerca. Revista de Pensament i Anàlisi*, n°4: 5-11
- González Esteban, E. (2013). Formación ética de los profesionales. Forjando el interés desde la razón y la emoción. *RIO. Revista Internacional de Organizaciones*, N° 10, Junio: 21-40.
- González Esteban, E. (2016). El desarrollo de la ética empresarial ante los avances de la neurociencia organizacional y la neuroética. *Pensamiento*, vol. 72, núm 273: 921-940.
- González Esteban, E. y Calvo Cabezas, P. (2019). Homo reciprocans from the neuroscientific literature: limited reciprocity. Criticism from neuroethics. *Anuario Filosófico* [revisión].
- Gould, S. J. (1997). The exaptive excellence of spandrels as a term and prototype. *Proceedings of the National Academy of Sciences USA*, 94: 10750-10755.
- Gould, S. J. (1999). *Rocks of Ages: Science and Religion in the Fullness of Life. The Library of Contemporary Thought*. New York: The Ballantine Publishing Group.
- Gould, S. J. y Lewontin, R. C. (1979). The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme. *The Royal Society of London, Series B, Vol. 205, N° 1161: 581-598*.
- Gould, S. J. y Vrba, E. S. (1982). Exaptation —A Missing Term in The Science of Form. *Paleobiology*, Vol. 8, N° 1: 4-15.

- Gracia, J. (2016). ¿Incorre la teoría del proceso dual de Joshua Greene en Falacia Naturalista?. *Pensamiento*, 72, 273: 789-807.
- Gracia, J. (2017). ¿Conduce la neurociencia a la naturalización del deontologismo?. *Pensamiento* 276: 581-586.
- Gracia, J. (2018). Crítica a la naturalización del deontologismo en la teoría del proceso dual del juicio moral de Joshua Greene. *Isegoría* n° 58.
- Grafman, J. et al. (1996). Frontal lobe injuries, violence, and aggression: a report of the Vietnam head injury study. *Neurology*, 46: 1231-1238.
- Greene J. D., R.B. Sommerville, L. E. Nystrom, J. M. Darley y J. D. Cohen. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, 293(5537): 2105-2108.
- Greene, J. D. (2003). From neural is to moral ought: what are the moral implications of neuroscientific moral psychology?. *Nature Reviews of Neuroscience*, vol. 4: 847-850.
- Greene, J. D. (2007). Why are VMPFC patients more utilitarian? A dual-process theory of moral judgment explains. *Trends in Cognitive Sciences*, Vol.11, N° 8: 322.
- Greene, J. D. (2008a). The Secret Joke of Kant's Soul. En Walter Sinnott-Armstrong (ed.), *Moral Psychology 3, The Neuroscience of Morality. Emotion, Brain Disorders, and Development*, 35-79. Cambridge, Massachusetts: MIT Press.
- Greene, J. D. (2008b). Reply to Mikhail and Timmons. En Sinnott-Armstrong, W. (ed.), *Moral Psychology 3, The Neuroscience of Morality. Emotion, Brain Disorders, and Development*, 105-117. Cambridge, MIT Press.
- Greene, J. D. (2009). Dual-process morality and the personal/impersonal distinction: A reply to McGuire, Langdon, Coltheart, and Mackenzie". En *Journal of Experimental Psychology*.
- Greene, J. D. (2013). *Moral Tribes. Emotion, Reason and the Gap between us and them*. New York: Pinguin.

- Greene, J. D. (2014a). Beyond Point-and-Shoot Morality: Why Cognitive (Neuro)Science Matters for Ethics. *Ethics*, 124: 695-726.
- Greene, J. D. (2014b). *The Cognitive Neuroscience of Moral Judgment and Decision Making*. MIT Press.
- Greene, J. D. (2015). The rise of moral cognition. *Cognition*, 135: 39-42.
- Greene, J. D. et al. (2004). The Neural Bases of Cognitive Conflict and Control in Moral Judgment. *Neuron*, 44: 389-400.
- Greene, J. D., F. A. Cushman, L. E. Stewart, K. Lowenberg, L. E. Nystrom y J. D. Cohen. (2009). Pushing moral buttons: The interaction between personal force and intention in moral judgment. *Cognition*, 111(3): 364-71.
- Greene, J. D., L. E. Nystrom, A. D. Engell, J. M. Darley, y J. D. Cohen. (2004). The Neural Bases of Cognitive Conflict and Control in Moral Judgment. *Neuron*, 44: 389-400.
- Greene, J. D., y J. Haidt. (2002). How (and where) does moral judgment work?. *Trends in Cognitive Science*, vol. 6, N° 12: 517-523.
- Habermas, J. (1999). *Conocimiento e interés*. Valencia: Universidad de Valencia.
- Habermas, J. (2002). *El futuro de la naturaleza humana: ¿hacia una eugenesia liberal?*. Barcelona: Paidós.
- Habermas, J. (2006). *Entre naturalismo y religión*. Barcelona: Paidós.
- Habermas, J. (2010). *Teoría de la acción comunicativa*. Madrid: Trotta.
- Haidt, J. (2001). The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment. *Psychological Review*, vol. 108, N° 4: 814-834.
- Haidt, J. (2013). *The righteous mind: why good people are divided by politics and religion*. London: Penguin Books.
- Hamilton, W. D. (1963). The Evolution of Altruistic Behavior. *American Naturalist*, 97: 354-356.

- Hamilton, W. D. (1964). The genetical evolution of social behaviour. *Journal of Theoretical Biology*, 7: 1-16.
- Harman, G. (1975). Reasons. *Critica*, 7: 3-13.
- Harman, G. (1977). *The Nature of Morality*. New York: Oxford University Press.
- Harman, O. (2012). Is the Naturalistic Fallacy Dead (and If So, Ought It Be?). *Journal of the History of Biology*, 45: 557-572
- Hauser, M. D. (2008). *La mente moral. Cómo la naturaleza ha desarrollado nuestro sentido del bien y del mal*. Barcelona: Paidós.
- Hauser, M. y Wood, J. (2010). Evolving the Capacity to Understand Actions, Intentions, and Goals. *Annual Review of Psychology*, 61: 303-324.
- Held, V. (2002). Moral Subjects: The natural and the normative. *Proceedings and Addresses of the American Philosophical Association*, 76(2): 7-24.
- Hill, J. (1976). *The Ethics of G. E. Moore: A New Interpretation*. Amherst, NY: Prometheus Books.
- Hobbes, T. (1999). *Leviatán: o la materia, forma y poder de un Estado eclesiástico y civil*. Madrid: Alianza.
- Holmes, O. W. (1897). The Path of the Law. *Harvard Law Review*, 10: 457-478.
- Honneth, A. (1997). *La lucha por el reconocimiento: por una gramática moral de los conflictos sociales*. Barcelona: Crítica.
- Hughes, W. (1986). Richard's Defence of Evolutionary Ethics. *Biology and Philosophy*, 1: 306-315.
- Hume, D. (1953 [1752]). *An Enquiry Concerning the Principles of Morals*. La Salle, IL: The Open Court Publishing, Co.
- Hume, D. (1988). *Tratado de la naturaleza humana*. Madrid: Tecnos.
- Huxley, T. H. (1894). *Evolution and ethics*. London: Macmillan and Co.

- Iacoboni, M. (2009). *Las neuronas espejo*. Barcelona: Katz.
- Illes, J. y Sahakian, B. J. (2011). *The Oxford Handbook of Neuroethics*. Oxford: Oxford University Press.
- Jácomo, A. (2011). Os contributos neurocientíficos actuais e a emergência da neuroética. *Revista Portuguesa de Bioética*, v. 57, Nº 15: 353-365.
- Jácomo, A. (2013a). E se o nosso eu fosse sináptico?. *Revista Portuguesa de Bioética*, v. 57, Nº 17: 103-134.
- Jácomo, A. (2013b). New bioethical challenges in neuroscience. *Bioethikos*, v. 7, Nº 1: 27-35.
- Jácomo, A. (2014). Existe neuroética?. *Actas das XVI jornadas Cofanor "Novidades do Cérebro"*. Porto.
- Jácomo, A. y D. Serrão. (2013) Implications of Neuroplasticity in Neuroethics. *Jerusalém Proceedings of the Jerusalem International Conference on Neural Plasticity and Cognitive Modifiability*, 73-78. Italia: Medimond.
- Jakovcevik, A., Irrazábal, M. y Bentosela, M. (2011). Cognición social en animales y humanos: ¿Es posible establecer un continuo?. *Suma Psicológica*, Vol. 18, Nº 1: 35-46.
- Johnston, M. (1992). How to Speak of the Colors. *Philosophical Studies*, 68: 221-263.
- Joyce, R. (2000). Darwinian Ethics And Error. *Biology and Philosophy*, 15: 713-732.
- Joyce, R. (2006). *The evolution of morality*. Cambridge, Massachusetts: MIT Press.
- Joyce, R. (2014). The origins of moral judgment. En Waal, F. et al. (eds.), *Evolved Morality. The Biology and Philosophy of Human Conscience*, 125-142). Leiden, Boston: BRILL.
- Kahane, G. y N. Shackel. (2010). Methodological Issues in the Neuroscience of Moral Judgment. *Mind and Language*, 25: 561-582.

- Kahane, G., K. Wiech, N. Shackel, M. Farias, J. Savulescu, y I. Tracey. (2012). The neural basis of intuitive and counterintuitive moral judgments. *Social Cognitive and Affective Neuroscience*, 7(4): 393-402.
- Kamm, F. M. (2009). Neuroscience and Moral Reasoning: A Note on Recent Research. *Philosophy and Public Affairs*, 37: 330-345.
- Kant, I. (1992). *Fundamentación de la Metafísica de las costumbres*. Madrid: Real Sociedad Matritense de Amigos del País.
- Kant, I. (2000). *Crítica de la razón práctica*. Madrid: Alianza.
- Kant, I. (1978). *Crítica de la razón pura*. Madrid: Alfaguara.
- Kitcher, P. (1985). *Vaulting Ambition: Sociobiology and the Quest for Human Nature*. Cambridge, Massachusetts: MIT Press.
- Kitcher, P. (1994). Four Ways of Biologizing Ethics. En E. Sober (ed.), *Conceptual Issues in Evolutionary Biology*. Cambridge, Massachusetts: MIT Press.
- Kohlberg, L. (1992). *Psicología del desarrollo moral*. Bilbao: Desclée de Browser.
- Kornhuber, H. H., Deecke, L., Lang, M., Lang, A. y Kornhuber, A. (1989). Will, volitional action, attention and cerebral potentials in man: Bereitschaftspotential, performance-related potentials, directed attention potential, EEG spectrum changes. En Wayne A. Hershberger (ed.), *Volitional action*, 107-168. Amsterdam: Elsevier.
- Kornhuber, H. H., y Deecke, L. (1965). Hirnpotentialänderungen bei Willkurbewegungen und passiven Bewegungen den Menschen: Bereitschaftspotential und reafferente Potentiale. *Pflugers Archiv für die gesamte Physiologie des Menschen und der Tiere*, 284(1): 1-17.
- Korsgaard, C. (2000). *Las fuentes de la normatividad*. México: Universidad Nacional Autónoma de México, Instituto de Investigaciones Filosóficas.
- Krause, S. (2008). *Civil Passions: Moral sentiment and democratic deliberation*. Princeton: Princeton University Press.

- Kurtz, P. (2007). What is the relationship among science, reason, and ethics?. En P. Kurtz (ed.), *Science and ethics: can science help us make wise moral judgments?*, 11-26. New York: Prometheus Books.
- Kurzban, R. et al. (2015). The Evolution of Altruism in Humans. *Annual Review of Psychology*. 66: 575-599.
- Lakatos, I. (2011). *Historia de la ciencia y sus reconstrucciones racionales*. Madrid: Tecnos.
- Le Doux, J. E. (1999). *El cerebro emocional*. Barcelona: Ariel.
- Lemos, J. (2000). Darwinian Natural Right and the Naturalistic Fallacy. *Biology and Philosophy*, 15: 119-132.
- Lenk, H. (1982). *Filosofía pragmática*. Barcelona: Alfa, Laia.
- Lenk, H. (1988). *Entre la epistemología y la ciencia social*. Barcelona: Alfa.
- Levy, N. (2007). *Neuroethics*. Nueva York: Cambridge University Press.
- Libet, B. (1999). Do We Have Free Will?. *Journal of Consciousness Studies*, Vol. 6, Nº 8-9, 1999: 47-57(11).
- Libet, B., Gleason, C. A., Wright, E. W. y Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness potential). *Brain*, 106: 623-642.
- Lorenzen, P. (1969). *Normative Logic and Ethics*. Mannheim: Bibliographisches Institut, Dudenverlag.
- MacIntyre, A. (2001). *Animales racionales y dependientes. Por qué los seres humanos necesitamos las virtudes*. Barcelona: Paidós.
- MacIntyre, A. C. (1959). Hume on "Is and "Ought". *The Philosophical Review*, 68(4): 451-468.
- Mackie, J. L. (1977). *Ethics: Inventing Right and Wrong*. Middlesex, England: Penguin Books.

- Mackie, J. L. (1980). *Hume's Moral Theory*. London: Routledge and Kegan Paul Ltd.
- Marazziti, D. et al. (2013). The neurobiology of moral sense: facts or hypotheses?. *Annals of General Psychiatry*, 12:6.
- Marcus, G. E. (2002). *The Sentimental Citizen: Emotion in Democratic Politics*. Pensilvania: Penn State Press.
- Maynard Smith, J. (1976). Evolution and the Theory of Games. *American Scientist*, 64: 41-55.
- Maynard Smith, J. y Price, G. R. (1973). The Logic of Animal Conflict. *Nature*, 246: 15-18.
- McCarthy, T. (1992). *Ideales e ilusiones*. Madrid: Tecnos.
- McShea, R. (1990). *Morality and Human Nature: A New Route to Ethical Theory*. Philadelphia: Temple University Press.
- McShea, R. y D. McShea. (1999). Biology and Value Theory. En M. Ruse y J. R. Maienschein, (eds.), *Biology and the Foundation of Ethics*, 307-327. New York: Cambridge University Press.
- Medina-Vicent, M. y Reverter Bañón, S. (2017). Monográfico Mujeres y Liderazgo. *Dossiers Feministes*, N. 22.
- Meltzoff, A. N. (2002). Imitation as a mechanism of social cognition: Origins of empathy, theory of mind, and the representation of action". En U. Goswami (ed.), *Handbook of childhood cognitive development*, 6-25. Oxford: Blackwell Publishers.
- Mihailov, E. (2015). The argument from self-defeating beliefs against deontology. *Ethical Perspectives*, 22(4): 573-600.
- Mihailov, E. (2016). Is Deontology a Moral Confabulation?. *Neuroethics*, 9: 1-13
- Mikhail, J. (2008). Comment on Greene. En W. Sinnott-Armstrong, *Moral Psychology 3, The Neuroscience of Morality. Emotion, Brain Disorders, and Development*, 80-91. Cambridge: MIT Press.

- Mill, J. S. (1968). *El utilitarismo*. Buenos Aires: Aguilar.
- Moll, J., P. J. Eslinger, y R. de Oliveira-Souza. (2001). Frontopolar and anterior temporal cortex activation in a moral judgment task: preliminary functional MRI results in normal subjects. *Arq. Neuropsiquiatr.* 59: 657-664.
- Moll, J., R. de Oliveira-Souza, I. E. Bramati, y J. Grafman. (2002a). Functional networks in emotional moral and nonmoral social judgments. *Neuroimage* 16: 696-703.
- Moll, J., R. de Oliveira-Souza, P. J. Eslinger, I. E. Bramati, J. Mourao-Miranda, P. A. Andreiuolo, y L. Pessoa. (2002b). The neural correlates of moral sensitivity: a functional magnetic resonance imaging investigation of basic and moral emotions. *The Journal of Neuroscience*, 22: 2730-2736.
- Moore, G. E. (1983). *Principia Ethica*. México: Universidad Nacional Autónoma de México.
- Mora, F. (2007). *Neurocultura: Una cultura basada en el cerebro*. Madrid: Alianza Editorial.
- Morgado, I. (2010). *Emociones e inteligencia social*. Barcelona: Ariel.
- Murphy, T. F. (2012). The Ethics of Impossible and Possible Changes to Human Nature. *Bioethics*, Vol. 26, N° 4: 191-197.
- Nagel, T. (1985). *The view from nowhere*. Oxford: Oxford University Press.
- Neurath, O. (1931). Physicalism: The Philosophy of the Viennese Circle. *The Monist*, 41, 4: 618-623.
- Nichols, S. (2005). Innateness and moral psychology. En Carruters, P., Laurence, S. y Stich, S. (eds.), *The innate mind: structure and contents*, 353-430. New York: Oxford University Press.
- Nussbaum, M. C. (2006). *Las fronteras de la justicia: Consideraciones sobre la exclusión*. Barcelona: Paidós.

- Ortega Esquembre, C. (2016). ¿Naturalizar la idea de justicia? Una respuesta crítica desde la teoría moral de Jürgen Habermas. *Pensamiento*, Vol. 72, Nº. 273, 827-848.
- Ortega Esquembre, C. (2018). ¿Naturalismo blando? En torno al estatuto epistemológico de la teoría moral de Jürgen Habermas. *Anuario Filosófico*, 51/1: 135-159.
- Ortiz de Landázuri, C. (2016). El error neurocientífico de Descartes, entre Spinoza y Tomás de Aquino. El debate entre Damasio y Stump sobre el materialism eliminativo en la neurociencia, neuropolítica y neuroeconomía. *Recerca*, 18, 107-133.
- Pallarés-Domínguez, D. (2013). Críticas y orientaciones para el estudio en neuroética. *Recerca*, 13, 85-102.
- Pallarés-Domínguez, D. (2016). My brain made me not do it: an emergentist interpretation of Benjamin Libet. *Ramon Llull Journal of Applied Ethics*, 7: 121-141.
- Pallarés-Domínguez, D. y González Esteban, E. (2016). The Ethical Implications of Considering Neurolaw as a New Power. *Ethics & Behavior*, vol.26, Issue 3: 252-266.
- Pallarés-Domínguez, D. y Richart, A. (eds.). (2018). Neuroética y neuroeducación: repensando la relación entre las neurociencias y las ciencias sociales. *Recerca*, 22.
- Paris Albert, S. y Comins Mingol, I. (2013). Epistemological and Anthropological Thoughts on Neurophilosophy: An Initial Framework. *Recerca*, 13: 63-88.
- Paxton, J. M., T. Bruni y J. D. Greene. (2014). Are 'counter-intuitive' deontological judgments really counter-intuitive? An empirical reply to Kahane et al. (2012). *Social Cognitive and Affective Neuroscience*, 9 (9): 1368-1371
- Pearson, J. M. (2014). Decision Making, The Neurological Turn. *Neuron*, 82: 950-965.
- Piaget, J. (1987). *El criterio moral en el niño*. Barcelona: Martínez Roca.

- Pierce, C. S. (1965). *Collected papers of Charles Sander Pierce*. Harvard: Harvard University Press.
- Pigliucci, M. (2003). On the relationship between science and ethics. *Zygon*, 38(4): 871-894.
- Popper, K. (1962). *La lógica de la investigación científica*. Madrid: Tecnos.
- Popper, K. (1967). *El desarrollo del conocimiento científico: conjeturas y refutaciones*. Buenos Aires: Paidós.
- Popper, K. (1992). *La Sociedad abierta y sus enemigos*. Barcelona: Paidós.
- Prinz, J. (2011). Where Do Morals Come From?. En Christen, M. (ed.). *Philosophical Implications of Empirically Informed Ethics*.
- Putnam, H. (1981). *Reason, Truth and History*. Cambridge: Cambridge University Press.
- Putnam, H. (1983). *Realism and Reason*. Cambridge: Cambridge University Press.
- Quine, W. V. O. (1953). Two Dogmas of Empiricism. En W. V. O. Quine, *From a Logical Point of View*. Cambridge: Harvard University Press.
- Quintelier, K., L. V. Speybroeck y J. Braeckman. (2011). Normative Ethics Does Not Need a Foundation: It Needs More Science. *Acta Biotheor.*, 59: 29-51.
- Radcliffe Richards, J. (2001). *Human Nature After Darwin: A philosophical introduction*. London: Routledge.
- Raine, A. et al. (2000). Reduced prefrontal gray matter volume and reduced autonomic activity in antisocial personality disorder. En *Arch. Gen. Psychiatry*, 57: 119-129.
- Rawls, J. (1995). *Teoría de la justicia*. México: Fondo de Cultura Económica.
- Rawls, J. (2013). *El liberalismo político*. Barcelona: Crítica.
- Regan, T. (1986). *Bloomsbury's Prophet*. Philadelphia: Temple University Press.
- Reverter Bañón, S. (2016). Reflexión crítica frente al neurosexismo. *Pensamiento*, 72(273): 959-979.

- Reverter Bañón, S. (2017). El Neurofeminismo frente a la investigación sobre la diferencia sexual. *Daimon. Revista Internacional de Filosofía* (Universidad de Murcia), suplemento 6: 95-110.
- Reverter Bañón, S. y Medina-Vicent, M. (2017). Intersecciones entre liderazgo y feminismo. *Dossiers Feministes*, N. 22: 5-12.
- Reverter Bañón, S. y Medina-Vicent, M. (2018). Neuroeducación para la democracia y la igualdad. A. Richart, M. García-Granero, C. Ortega Esquembre y L. de Tienda Palop (Eds.), *Juicio moral y democracia. retos de la ética y la filosofía política*, 314-325. Granada: Comares.
- Richards, R. J. (1986a). Justification through biological faith: a rejoinder. *Biology and Philosophy*, 1(1): 337-354.
- Richards, R. J. (1986b). A Defence of Evolutionary Ethics. *Biology and Philosophy*, 1: 265-293.
- Richards, R. J. (1987). *Darwin and the emergence of evolutionary theories of mind and behaviour*. Chicago: University of Chicago Press.
- Richart, A. (2016). El origen evolutivo de la agencia moral y sus implicaciones para la ética. *Pensamiento*, Vol. 72, Nº. 273, 849-864.
- Rizzolatti, G. y Sinigaglia, C. (2006). *Las neuronas espejo. Los mecanismos de la empatía emocional*. Barcelona: Paidós.
- Rorty, R. (2004). The Brain as Hardware, Culture as Software. *Inquiry*, 47 (3): 219-235.
- Rosenberg, A. (2000). *Darwinism in philosophy, Social Science and Policy*. Cambridge: Cambridge University Press.
- Roskies, A. L. (2010). How Does Neuroscience Affect Our Conception of Volition?. *Annual Review of Neuroscience*, 33: 109-130.
- Roskies, A. L. (2010). Why Libet's studies don't pose a threat to free will. En Sinnott-Armstrong, W. y Nadel, L. (eds.), *Conscious Will and Responsibility*, 11-22. New York: Oxford University Press.

- Rottschaefer, W. A. (1997). The Scientific Naturalization of Ethics. *ResearchGate*: <https://www.researchgate.net/publication/242582486> The Scientific Naturalization of Ethics
- Rottschaefer, W. A. (1998). *The Biology and Psychology of Moral Agency*. Cambridge: Cambridge University Press.
- Rottschaefer, W. A. (2007). Scientific naturalistic ethics. Weird science and pseudoethics?. En P. Kurtz y D. Koepsell (eds.), *Science and ethics. Can science help us make wise moral judgments?*, 285-305. New York: Prometheus Books.
- Rottschaefer, W. A. y D. Martinsen. (1990). Really Taking Darwin Seriously: An Alternative to Michael Ruse's Darwinian Metaethics. *Biology and Philosophy*, 5: 149-173.
- Ruse, M. (1979). *Sociobiology: Sense or Nonsense?*. Dordrecht, Holland: D. Reidel Publishing Co.
- Ruse, M. (1982). *Darwinism Defended: A Guide to the Evolution Controversies*. Reading, Massachusetts: Addison-Wesley.
- Ruse, M. (1986a). Evolutionary Ethics: A Phoenix Arisen. *Zygon*, 21(1): 95-112.
- Ruse, M. (1986b). *Taking Darwin Seriously*. Oxford: Basil Blackwell.
- Ruse, M. (1995). *Evolutionary Naturalism: Selected essays*. London: Routledge.
- Ruse, M. (2008). *Charles Darwin (Vol. 5)*. Oxford: Blackwell.
- Ruse, M. y E. O. Wilson. (1986). Moral philosophy as applied science. *Philosophy*, 61: 173-192.
- Russell, B. (1966). *Ensayos sobre lógica y conocimiento*. Madrid: Taurus.
- Salles, A. (2013). On the normative implications of social neuroscience. *Recerca*, 13: 29-42.
- Salles, A. (2014). Neuroethics in a "Psy" World: the Case of Argentina. *Cambridge Quarterly of Healthcare Ethics*, 23:3.

- Salles, A. y Evers, K. (2014). *La vida social del cerebro*. México: Editorial Fontamara.
- Salles, A. y Evers, K. (2017). Social Neuroscience and Neuroethics: A Fruitful Synergy. En A. Ibáñez et al. (eds.) *Neuroscience and Social Science: The Missing Link*, 531-546. Springer.
- Salles, A. (2015). Rationality and the moral significance of emotions. *Inherent and Instrumental Values: Excursions in Value Inquiry*, 89-99. University Press of America.
- Salles, A. (2016). Sobre la neuroética. *Revista Latinoamericana de Filosofía*, vol. 42.
- Salles, A. (2017). Proactive Epigenesis and Ethics. *EMBO Reports*, vol. 18, N° 8 ss. 1271.
- Salles, A. (2018). Neuroethics in Context: The Development of the Discipline in Argentina. En Johnson, S. y Rommelfanger, K. (ed.) *The Routledge Handbook of Neuroethics*. Routledge.
- Salles, A. Brain Imaging and Privacy Concerns (2016). M. Farisco y K. Evers (eds.) *Neurotechnology and Direct Brain Communication*. New York: Routledge.
- Salles, A., Farisco, M. y Evers, K. (forthcoming) Neuroethics and Philosophy in Responsible Research and Innovation: the case of the HBP. *Neuroethics*.
- Schwemmer, O. (1980). *Philosophie der Praxis*. Frankfurt am Main: Suhrkamp Verlag.
- Searle, J. (1992). *Intencionalidad*. Madrid: Tecnos.
- Searle, J. (1996). *El redescubrimiento de la mente*. Barcelona: Crítica.
- Searle, J. (2006). *La mente, una breve introducción*. Bogotá: Norma.
- Sen, A. (1979). *La desigualdad económica*. Barcelona: Crítica.
- Serrallonga, J. (2005). No estamos solos: australopitecos y chimpancés habilidosos. En Guillén-Salazar, F. (ed.), *Existo, luego pienso: Los Primates y la Evolución de la Inteligencia Humana*, 171-251. Madrid: Ateles.

- Seymour, B. y Dolan, R. (2008). Emotion, Decision Making, and the Amygdala. *Neuron*, 58: 662-671.
- Shenhav, A. y Greene, J. D. (2010). Moral Judgments Recruit Domain-General Valuation Mechanisms to Integrate Representations of Probability and Magnitude. *Neuron*, 67: 667-677.
- Sherman, N. (1999). Taking Responsibility for our Emotions. *Social Philosophy and Policy*, Vol. 16, 2: 294-323.
- Singer, P. (1981). *The expanding circle: Ethics and sociobiology*. New York: Farrar Straus & Giroux.
- Singer, P. (2011). *The expanding circle: Ethics, Evolution and Moral Progress*. Princeton: Princeton University Press.
- Singer, W. (2004). Selbsterfahrung und neurobiologische Fremdbeschreibung. *Deutsche Zeitschrift für Philosophie*, LII, 8: 235-256.
- Siurana, J. C. (2011a). Las capacidades ético-discursivas como capacidades para el reconocimiento recíproco: bases de un modelo de desarrollo humano. *Éthique et économique/Ethics and Economics*, 8 (1).
- Siurana, J. C. (2011b). *Los consejos de los filósofos: una introducción a la historia de la ética*. Barcelona: Proteus.
- Siurana, J. C. (2015). *Ética del humor: fundamentos y aplicaciones de una nueva teoría ética*. Madrid: Plaza y Valdés.
- Siurana, J. C. (2017). *Felicidad a golpe de autoayuda. Tu vida en manos de un best seller*. Madrid: Plaza y Valdés.
- Skinner, B. F. (1977). *Sobre el conductismo*. México: Trillas.
- Sloan Wilson, D., E. Dietrich y A. B. Clark. (2003). On the inappropriate use of the naturalistic fallacy in evolutionary psychology. *Biology and Philosophy*, 18: 669-682.

- Sober, E. (1994). *From a Biological Point of View*. Cambridge, UK: Cambridge University Press.
- Spencer, H. (1864). *The principles of Biology*. London: Williams and Norgate.
- Spencer, H. (1983). *The principles of sociology*. London: Williams and Norgate.
- Sripada, C. y Stich, S. (2005). A framework for the psychology of norms. En P. Carruthers, S. Laurence, y S. Stich (eds.), *The innate mind: Structure and content*, 280-301. New York, NY: Oxford University Press.
- Stuchlick, J. (2011). Felicitology: Neurath's Naturalization of Ethics. *Hopos*, 1(2): 183-208.
- Sturgeon, N. L. (2003). Moore on Ethical Naturalism. *Ethics*, 113: 528-556.
- Tagney, P. et al. (2007). Moral Emotions and Moral Behavior. *Annual Review of Psychology*, 58: 345-572.
- Teehan, J. y C. diCarlo. (2004). On the Naturalistic Fallacy: A Conceptual Basis for Evolutionary Ethics. *Evolutionary Psychology*, 2: 32-46.
- Thornhill, R., y Palmer, C. T. (2000). *Rape: a natural history*. Cambridge: MIT Press.
- Timmons, T. (2008). Towards a Sentimentalist Deontology. En Sinnott-Armstrong, W., *Moral Psychology 3, The Neuroscience of Morality. Emotion, Brain Disorders, and Development*, 93-104. Cambridge: MIT Press.
- Tomasello, M. y Vaish, A. (2013). Origins of Human Morality. *Annual Review of Psychology*, 64: 231-255.
- Torralva, T., y Manes, F. (2014). Cognición social. En Salles, A. y Evers, K. (coords.), *La vida social del cerebro*, 17-42. México: Fontamara.
- Trivers, R. L. (1971). The Evolution of Reciprocal Altruism. *The Quarterly Review of Biology*, 46: 35-57.
- Trivers, R. L. (1985). *Social evolution*. California: Benjamin Cummings.

- Waal, F. (1997). *Bien natural: Los orígenes del bien y del mal en humanos y otros animales*. Barcelona: Herder.
- Waal, F. (2007). *Primates y filósofos: La evolución de la moral del simio al hombre*. Barcelona: Paidós.
- Walsh, F. M. (2008). The return of the naturalistic fallacy: a dialogue on human flourishing. *Heythrop J.*, 49(3): 370-387.
- Walter, A. (2006). The anti-naturalistic fallacy: Evolutionary moral psychology and the insistence of brute facts. *Evolutionary Psychology*, 4: 33-48.
- Warnock, M. (1978). *Ethics Since 1900*. Oxford: Oxford University Press.
- Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological Review*, 20: 158-177.
- Weston, A. (2006). *A Practical Companion to Ethics*. Oxford: Oxford University Press.
- Wielenberg, E. J. (2014). Review of Greene's Moral Tribes. *Ethics* 124 (4): 910-916.
- Williams, B. (1972). *Morality*. Cambridge: Cambridge University Press.
- Williams, G. C. (1966). *Adaptation and natural selection*. Princeton University Press: Princeton, New Jersey.
- Wilson, D. S., y Sober, E. (1994). Reintroducing Group Selection to the Human Behavioral Sciences. *Behavioral & Brain Sciences*, 17: 585-654.
- Wilson, E. O. (1975). *Sociobiology: The New Synthesis*. Cambridge: Harvard University Press.
- Wilson, E. O. (1984). *Biophilia*. Harvard: Harvard University Press.
- Wilson, E. O. (1998). *Consilience*. New York: Knopf.
- Winch, P. (2003). *The Idea of a Social Science: and its Relationship with Philosophy*. London: Routledge.

Wittgenstein, L. (1988). *Investigaciones filosóficas*. México: Instituto de investigaciones filosóficas.

Woolcock, P. (1993). Ruse's Darwinian Meta-Ethics: A Critique. *Biology and Philosophy*, 8: 423-439.

Woolcock, P. (1999). The case against evolutionary ethics today. En J. R. Maienschein y M. Ruse, (eds.), *Biology and the foundation of ethics*, 276-306. Cambridge: Cambridge University Press.

Wright, R. (1994). *The Moral Animal: The new science of evolutionary psychology*. New York: Pantheon Books.

Wynee-Edwards, V. C. (1962). *Animal Dispersion in Relation to Social Behaviour*. Edimburgo: Oliver and Boyd.