

Universitat de València

Facultat de Filosofia i Ciències de l'Educació

**Departament de Filosofia, Metafísica i Teoria del Coneixement, Lògica i
Filosofia de la Ciència i Filosofia Moral**



VNIVERSITAT DE VALÈNCIA

**LOS SUPUESTOS METAFÍSICOS DEL PROBLEMA DIFÍCIL DE LA
EXPERIENCIA CONSCIENTE EN LA FILOSOFÍA DE DAVID
CHALMERS**

Tesis Doctoral

**Presentada al título de Doctor en Filosofía,
con Mención Internacional**

Doctorando: Alejandro Fontcuberta Llavata

Director: Josep E. Corbí Fernández de Ibarra

Programa de Doctorado: Pensamiento Filosófico Contemporáneo

Octubre de 2016

A la memoria de mi madre y de mi abuela

ÍNDICE

Agradecimientos	vii
Précis (Resumen)	x
Introduction (Introducción)	xxxiii
1. Un marco general	1
1.1. Fisicismo y supervinencia	2
1.2. Fisicismo y microfisicismo	9
1.3. Explicación reductiva e iluminación	12
1.4. La geografía de la discusión	17
1.5. La concepción absoluta de la realidad, explicaciones desenmascaradoras y la noción de “dar sentido”	23
2. Los dos conceptos de lo mental: una distinción problemática	36
2.1. Una primera aproximación.....	36
2.2. El catálogo y los dos conceptos de lo mental.....	42
2.3. La doble vida de los estados mentales y la co-ocurrencia de las propiedades psicológicas y fenoménicas	48
2.4. El problema de la individuación de las cualidades fenoménicas	51
3. Argumentos a favor de la irreductibilidad de la conciencia	60
3.1 Los argumentos	61
3.1.1. La posibilidad lógica de los zombis.....	61
3.1.2. El espectro invertido	66
3.1.3. Asimetría epistémica.....	67
3.1.4. El argumento del conocimiento	68
3.1.5. La ausencia de análisis	68
3.2. Consideraciones generales sobre los argumentos.....	69
3.3. Concebibilidad, posibilidad lógica y necesidad	80
4. Epifenomenismo y experiencia consciente	94
4.1. El epifenomenismo.....	95
4.1.1. El fisicismo y la eficacia causal de lo mental	97
4.1.2. Supervinencia, funcionalismo y causalidad.....	100
4.1.3. La concepción estratificada de la realidad y el drenaje de los poderes causales	103
4.1.4. Propiedades intrínsecas	106
4.2. La paradoja del juicio fenoménico	120
4.2.1. Los tipos de juicios fenoménicos.....	121

4.2.2. La paradoja.....	124
4.2.3. Experiencia consciente, relevancia causal y familiaridad	125
4.3. Epifenomenismo y atribuciones de experiencia consciente.....	131
5. Experiencia consciente, realismo y el Teatro Cartesiano.....	136
5.1. Experiencia consciente, apariencia y realidad	137
5.2. Realismo y familiaridad	139
5.3. El materialismo cartesiano	142
5.4. La crítica del Teatro Cartesiano.....	150
5.4.1. Revisiones estalinistas y revisiones orwellianas	150
5.4.2. El Teatro Cartesiano y el verificacionismo	155
5.4.3. Escepticismo y sustantividad.....	156
5.4.4. La Ilusión de Capgras.....	159
6. Experiencia Consciente y Valor	165
6.1. Introducción	165
6.2. El Estudio de la mente y lo inconsciente.....	167
6.2.1. Visión ciega.....	170
6.2.2. Lo inconsciente y lo cotidiano	173
6.3. Autoconocimiento y conciencia	179
6.3.1. El detectivismo	181
6.3.2. “Consciente de” y “conscientemente”	187
6.4. Experiencia, pasividad y acción	190
6.5. La ciencia de la conciencia	198
6.5.1. La ciencia de la conciencia	199
6.5.2. Dualismo naturalista y fisicismo.....	204
6.5.3. Propiedades fenoménicas	208
6.5.4. Ciencia y exhaustividad	217
Conclusions (Conclusiones)	222
Bibliografía.....	230

Agradecimientos

La historia de cómo se gestó esta tesis doctoral no tiene una fecha clara de inicio. Fue en 2002 cuando, de la mano de Josep E. Corbí, tuve el primer contacto con la filosofía de la mente, y desde ese momento mi interés apasionado por todos los temas que conforman dicha área no solo no ha disminuido, sino que se ha enriquecido y ramificado a otras áreas de la reflexión filosófica. Lo que me dejó esta primera experiencia con la filosofía de la mente no solo fue una base sólida que más adelante me permitió profundizar en los temas que más me interesaban de un modo autónomo, sino que me convencí de que la filosofía sí nos puede ayudar a comprender mejor el mundo, y que el tipo de comprensión que nos ofrece difícilmente podría provenir de otras disciplinas académicas.

Creo que es habitual entre los estudiantes de filosofía llegar a un punto en el que el fuerte escepticismo generado por discusiones que nunca tienen un claro vencedor acaba provocando una apatía, quizá irreversible, hacia la reflexión filosófica o, en ocasiones, un cinismo que, para muchos de los que continúan dedicándose a la filosofía profesionalmente, se disfraza de oscurantismo o de una aparente adaptación satisfactoria a las corrientes que configuran el terreno de lo aceptable. En estos casos, hacer filosofía no debe de ser algo muy distinto a desempeñarse en cualquier actividad con el objetivo de sobrevivir. Pero si este fuera el único *pathos* que mueve al filósofo profesional, diría que la filosofía no solo es prescindible, sino que hay muchas otras actividades que permiten la supervivencia y que seguramente son mucho más saludables tanto para el cuerpo como para la mente. Creo que la única manera de entender lo que puede aportarnos la filosofía pasa por comprender que la filosofía es un camino íntimo, que solo uno puede recorrer y que solo tiene sentido cuando se percibe que transitar dicho camino responde a una necesidad humana. Por otra parte, creo que es muy difícil lograr esta convicción sin una guía adecuada. Me siento afortunado de poder decir que, durante toda mi etapa de estudiante, no solo he podido disponer de dicha guía, sino que constantemente me he sentido motivado y estimulado para seguir

profundizando en la comprensión de preguntas que me acompañan constantemente.

Quiero agradecer especialmente a Josep el apoyo que me ha dado todos estos años, no solo por mostrarme lo valiosa que era la filosofía en mi vida, sino como un amigo que siempre estuvo presente, incluso en momentos muy difíciles para mí. Sin duda, esta tesis habría sido muy diferente y peor sin el concienzudo trabajo de Josep, quien leyó varias versiones completas de la tesis y contribuyó a su mejora con un detallado listado de comentarios y observaciones. Agradezco también el haber podido disfrutar de un ambiente intelectual saludable y lleno de pasión por la filosofía dentro del que fuera el *Departament de Metafísica i Teoria del Coneixement*. Quiero dar las gracias, especialmente, a Tobies Grimaltos, Carlos Moya y Julián Marrades; con ellos no solo aprendí incontables cosas, sino que siempre fueron un ejemplo de honestidad intelectual. Asimismo, no puedo dejar de referirme a los miembros (o ex miembros) del grupo *Phrónesis* de Valencia: Claudia Compte, Josep E. Corbí, Manolo Garcés, Miracle Garrido, Tobies Grimaltos, Carlos Moya, Eduardo Ortiz, José Luis Pérez, Lino San Juan, Sergi Rosell y Jordi Valor, con quienes he compartido con entusiasmo horas de discusiones en numerosos seminarios. También agradezco las oportunidades que he tenido de compartir espacios de discusión con Manuel García-Carpintero, Josep Lluís Prades, José Martínez, Jennifer Saul y Gianfranco Soldati. Cito también los nombres de miembros y antiguos miembros de departamento con los que tuve la oportunidad de discutir y aprender mucho, más de lo que puedo llegar a reconocer en estos agradecimientos, me refiero a Noemí Calabuig, Miriam Hoyo, Joan B. Llinares, Julián Marrades, Loris Pasinato, Vicente Raga y Vicente Sanfélix, y también quiero recordar a mis grandes amigos, todos, de algún modo u otro, filósofos: Boris Adamovich, Xisela García, Rafael Lidó, Jorge Magdaleno, Juan Carlos Mosquera, Paula Sáez, Gonzalo Soria, Carlos Tomás y Olga Varela. Una mención especial merece Mari Mikkola, quien me dio la bienvenida muy amablemente en sus cursos y seminarios impartidos en la Humboldt Universität zu Berlin, y en los que aprendí a reflexionar más allá de límites que ni sabía que existían.

Finalmente, quiero agradecerle a mi familia todo el apoyo que me ha prestado durante todo este tiempo, aunque haya sido en gran parte desde la distancia. Le doy las gracias especialmente a mi hermano, a mi padre, a María, a Clara y a Sergio, y doy gracias a la vida por poder sentirme acompañado por mi pequeña hija Gabriela, quien, sin saberlo, dio grandes empujones a la elaboración final de este trabajo con su sonrisa. No puedo olvidar a mi abuela, quien lamentablemente falleció unos meses antes de la finalización de este trabajo. A la memoria de ella y de mi madre, dedico esta tesis.

Bogotá, septiembre de 2016

Précis (Resumen)

Chapter 1 of this dissertation deals with the physicalist framework that Chalmers takes as a starting point. No specific definition of “physical property” plays any special role in the characterization of physicalism as a philosophical stance. Thus, for the sake of argument I will take for granted that “physical property” refers to the kind of properties current physics refers to and to the kind of intelligibility physics provides of natural phenomena. Chalmers articulates physicalism appealing to the notion of supervenience:

B-properties supervene on A-properties if no two possible situations are indiscernible with respect to their A-facts while differing in their B-facts. (Chalmers 1996: 33)

According to this scheme, there can be different types of supervenience depending on whether we refer to individuals or worlds, local or global supervenience, in each case, or whether we understand by possibility logical or natural possibility, so that there is logical and natural supervenience, respectively. According to Chalmers, it is logical supervenience the relevant notion to define physicalism, so that “materialism is true if all the positive facts of the world are globally logically supervenient on the physical facts” (Chalmers 1996: 41).

The problem with this definition is that it is still compatible with paradigmatically non-physicalist positions, like some forms of dualism. Supervenience establishes a brute correlation but it does not explain why this relation holds. On the other hand, Identity between mental and physical properties is a too strong relation because it excludes the possibility that mental properties are exemplified by different kinds of material. According to Chalmers, what characterizes a property apt to be reduced to physical properties is that it plays a certain causal role, regardless of how this property is realized. Explaining reductively certain property amounts to analysing it functionally and discovering empirically how it is realized; and this kind of explanation would make it intelligible the truth of physicalism. This position in philosophy of mind is called functionalism. Chalmers

argues that every mental phenomenon can be functionally analysable except consciousness.

Chalmers tends to speak interchangeably about physical properties and microphysical properties, but physicalism and microphysicalism are different theses. Microphysicalism is committed to a hierarchical layered reality, where properties of a determinate level are realized by properties of an inferior level. Additionally, microphysical properties are the fundamental properties in this hierarchical model. A non-microphysicalist physicalist is free of commitments with this hierarchy; but Chalmers conception of physicalism and its intuitive strength seems to rely on this hierarchical picture, rather than on a definite conception of physical.

Then, according to Chalmers a phenomenon is reductively explainable if its instantiation is globally logically supervenient on physical properties. Chalmers recognises that this definition could be too weak for two reasons: firstly, the facts to which the phenomenon is reduced could be an arbitrary set of details without any explanatory unit; secondly, given that the phenomenon in question could be realized by disparate realizers, this explanation cannot account for the phenomenon as a type. But at least, this kind of explanation would be a necessary condition of any reductive explanation and its failure would mean the failure of reductive explanation tout court. This explanation would be, additionally, what Chalmers calls a mystery-removing explanation: an explanation that reduces the bruteness and arbitrariness of a phenomenon to the bruteness and arbitrariness of lower-level processes. This explanation, though, is not an illuminating one, according to Chalmers, even though no further definition of illuminating explanation is given by Chalmers.

Chalmers' position is composed by two statements: first, physicalism is false; second, conscious experience supervenes either logically or naturally on physical facts. Given that physicalism is false, because, as we will see, facts about conscious experience wouldn't supervene logically on physical facts, then consciousness supervenes naturally on physical facts. Chalmers' arguments, as we will see later, rely on the idea that there is an epistemic gap between consciousness and physical facts, and this epistemic gap allows us to infer that there is an ontological gap between consciousness and physical facts, so that materialism is false. According

to Chalmers, the spectrum of answers to his antimaterialist arguments reduces either to denying the existence of an epistemic gap, or to denying the inference from an epistemic gap to an ontological one. There is a third option according to which there is an epistemic gap that will be closed in the future. The alternative approach to these types of materialism is embracing some or another form of dualism, or else a monism called by Chalmers Russellian monism, which has things in common with dualism.

It is not clear the relation between physicalism and the different perspectives hailing from the cognitive peculiarities of the cognitive beings we are, and this is an issue Chalmers does not address. These peculiarities could be subsumed in the privileged description of physics, or else be explained away as prejudices and biases that do not represent the world as it really is. This kind of project is what Bernard Williams called the absolute conception of reality. Of course, to carry out this project, we have to take as a starting our conceptual system. But it is evident that some of our concepts have an arbitrary source that can be cultural, historical or even biological, so that they don't correspond to reality as it is independently of us. In any case, in order to begin with this project, our starting point must be the concepts we have at our disposal. However, according to Williams, there can be some special concepts and beliefs that enjoy a privileged status due to its explanatory power, namely, the concepts of modern science. The history of modern science is full with examples where a widespread theory or conception is rejected in the light of the explanatory power of the new theories. And even though physicalism does not overlap with the project of an absolute conception of reality it shares with it, on the one hand, the aspiration of becoming a privileged description of reality; on the other hand, both are fuelled by the explanatory power of modern science and, more concretely, of physics.

However, what is especial about the explanatory role of physical properties? What makes them especial is that the set of causal networks in which they are embedded are independent of the mind, insofar as they exist independently of the existence of perceptual beings that come to have perceptions in virtue of the effect of some physical properties. However, the concept of causation used in this discussion is a commonsensical and pretheoretical one (not a physical one), and our

everyday concept of causation is tied to our regional perspectives, so that it is not clear that it could serve the purposes of an absolute conception of reality.

It is undeniable that the success of modern science is linked to the inhibition of some subjective elements that could bias the scientific research. Ideally, no matter who is conducting a scientific research, whenever she respects some methodological rules that help us to have an objective knowledge of reality. The correct application of these rules has allowed us rejecting mistaken conceptions of reality. This kind of explanation where something is explained away without presupposing its truth is what Barry Stroud calls an unmasking explanation. Physicalism tries to show that mental events are further instances of physical things, unmasking its true nature. The problem is how physicalism can give content to its explanandum starting from the austere language of physics. On the other hand, an unmasking explanation regarding the mind should account for the relation between physical and mental facts. However, Chalmers approach only recognises a brute correlation between both kinds of facts, but this model of explanation cannot give a proper understanding of how our perceptions relate to the world.

Williams distinguishes a kind of reflection that, not being independent of producing a true speech, is not limited to saying something true. That's the notion of "making sense". Making sense of something is always linked to a particular narrative and only makes sense to a particular audience. What makes sense to some group does not make sense to another one. The notion of "making sense" cannot be linked to an exclusive or privileged narrative, without taking into account a particular audience, and even science is a kind of narrative. A key question to be posed about physicalism is to what extent the kind of narrative constituted by physicalism can account for all the aspects that are relevant for human beings and especially regarding consciousness, that is allegedly what makes life worth living. In Chapter 6, this subject will be dealt with.

Chapter 2 aims to clarify the concept of consciousness and focuses on a distinction Chalmers draws to pin down a genuinely philosophical problem, in contrast to a series of problems that are amenable to a solution from the field of cognitive science. These are the basic questions Chalmers poses: "How could a

physical system such as a brain also be an *experiencer*? Why should there be *something it is like* to be such a system?" (Chalmers 1996: xi). These questions point to the qualitative aspect of the experience, which is not to be conflated with the causal processes and the processing of information accompanying this qualitative field. This distinction is illustrated by means of different metaphors that do not fix a clear concept (like the opposition between external and internal).

On the other hand, there are different characteristics that initially could help us to identify a philosophically relevant concept of consciousness. Our mental states can only be had by us; but this is not an exclusive feature of the mind, the history of a determinate country can only be had by that country. Another identity criterion could be the scepticism we can always have about the mental states of people different from ours, but sceptical doubts are not exclusive to the mind either. However, in the case of consciousness, it seems that we hold a special relation to our own mental states and we are debarred from getting to know what a very different form of life is like to be. This latter characterization is inseparable from the formulation of a problem, being a circular definition, unless consciousness is a primitive notion and cannot be defined beyond the illustration of some examples.

Chalmers counts amongst mental conscious states the following types: visual experiences, auditory experiences, tactile experiences, olfactory experiences, taste experiences, experiences of hot and cold, pain, bodily sensations, mental imagery, conscious thought, emotions, the sense of the self and some additional experiences. Nonetheless, it is not clear what these items have in common with each other. It can be said that for each of the items there are researches in the field of cognitive science. But what is the object of study of cognitive science? Saying that it is the mind will be of no avail, insofar as there can be different conceptions of the mind. On the other hand, unconscious states are mental states too.

Chalmers distinguishes between the psychological concept of consciousness and the phenomenal one. The psychological concept is understood as the causal or explanatory basis for behaviour, so that a mental state is such if it plays the appropriate causal role. This concept is the object of study of cognitive science and Chalmers refers to it as awareness. Phenomenal consciousness, on the other hand,

is conscious experience, understood as the set of consciously experienced mental states. According to Chalmers, both concepts exhaust the mental. The problems linked to the psychological concept are the so called easy problems: even though they are still to be solved, they do not pose any philosophical enigma. On the contrary, phenomenal consciousness raises the problem of why phenomenology exists, why cognitive functioning is accompanied by conscious experience, and this is the hard problem of consciousness.

Is there a psychological and phenomenal aspect for each mental state? According to Chalmers, nothing important rests on whether one of the aspects is essential to consider an instance of one state as belonging to one or another kind. Everyday mental concepts combine both aspects. The only important thing is to be clear regarding the aspect we are taking into account in each particular case. These two concepts are not distinguished in everyday practices, but this is, Chalmers says, because they empirically co-occur; and even if the extension of both concepts is the same, they are conceptually distinguishable. In any case, Chalmers warns us against the error of conflating both aspects of the mental, even if the criterion that allow us to fix the reference of phenomenology is the psychological aspect.

The problem that Chalmers has to face is that of individuating phenomenal mental states independently of this psychological aspect. Given that the phenomenal and the psychological aspects of the mind hold an empirical relation, their identity conditions must be independent of that relation; otherwise, that relation is not contingent, but logically necessary. Initially, psychological processes seem to have clear identity conditions. They are identical to certain causal roles. The question is why choosing a determinate causal profile as the one to be identified with the relevant mental state; how can we be sure that a determinate causal pattern is accompanied by phenomenology? How can we choose a set of causal relations as the relevant mental state, without taking into account phenomenology? Otherwise, how do we determine the start and the end, for instance, of a pain behaviour? It seems, following Chalmers distinction, that phenomenology is logically prior. But being this so, instances of phenomenal mental properties should be individuated before we are able to identify a determinate causal profile.

According to Chalmers, the only way we have to know about phenomenology is from the first person point of view. But in this case, first person private evidence introduces a huge difficulty. Following Peter F. Strawson's argument in *Individuals* (1957), there are some basic particulars that allow us to refer to other kinds of particulars. Strawson identified physical bodies and persons as such basic particulars, so that, for example, an instance of pain must be the pain of a particular body. Our body occupies a special causal role in relation to our own mental states. But why our conscious mental states must be attributed to something? And why do we attribute conscious mental states to the same things to which we attribute physical characteristics? One possible answer to the second question is that each kind of mental state is attributed to a different substance; this is Descartes' dualism. I consider this kind of dualism is too implausible. According to another kind of answer, even the first question would be mistaken, insofar as mental states are not attributed to any subject. The special causal role of bodies leads us to think mistakenly that mental states are possessed by a subject, there is a causal relation between mental states and the body, and that's all there is to having an experience. The problem with this latter position is that something akin to the following statement could be formulated coherently: "All *my* experiences are had by the body B". But what does it mean "my" in that sentence? "My" cannot mean the experiences that hold a special relation to the body B, because that would be a tautology, and the relation between experiences and body should be contingent, according to the previous reasoning. In order to individuate the set of "my own" experiences, I need to identify it among the experiences of others, so that a precondition to self-ascribe conscious mental states is to be able to ascribe mental states to others. However, if this argument is right, private phenomenology cannot serve as a starting point to individuate mental states and Chalmers distinction between phenomenal consciousness and psychology cannot be drawn intelligibly. A further consequence of this argument is that the relation between our conscious mental states and public criteria of identification cannot be merely empirical, as Chalmers suggests, but logical.

Chapter 3 deals with the antimaterialist arguments regarding phenomenal consciousness. The arguments can be sorted according to three criteria. First, there are arguments that start from the conception of a counterfactual situation and draw conclusions about what is logically possible. The metaphysical possibility of zombies is the most important argument of this type. Zombies are functionally and physically like us, but they lack phenomenal conscious experience. The basic form of the argument is as follows:

- (1) Zombies are conceivable.
- (2) If zombies are conceivable, then they are logically possible.
- (3) If zombies are logically possible, then materialism is false.
- (4) Materialism is false.

The inverted spectrum is the second argument of this type. This argument, as Chalmers formulates it, takes as a starting point the possibility that two physical twins have their colour experiences inverted, so that when one of them sees red, for example, the other one sees blue and vice versa. If this is a conceivable possibility, could we fix a set of physical facts that determine chromatic experiences? Following the schema of the previous argument, if the answer to this question is negative, then conceivability would entail logical possibility and materialism would be false.

The second type of arguments has to do with the asymmetry between how we relate to our own mental states and how we relate to the mental states of others. It seems that we have some kind of privilege when it comes to know our own mental states, and allegedly we cannot be wrong about them. But when it comes to detect mental states from a third person point of view, we could be systematically wrong, even in ideal epistemic situations. According to Chalmers, from these epistemic premises a metaphysical conclusion can be drawn, namely, that consciousness is irreducible. An argument that resorts to the same kind of intuition is Frank Jackson's

knowledge argument. This argument explores a conceptual experiment where Mary, a neuroscientist that has been raised in a colourless environment, knows every neurological and physical fact about colour. However, in spite of this knowledge, Mary does not know what colours look like, so if she were shown a coloured thing she would learn something new; but that would mean that physical facts do not exhaust all the facts and physicalism would be false.

The last kind of argument is based on the absence of a functional analysis of phenomenal consciousness. Phenomenal consciousness, by definition, is not conceptually linked to a determinate causal role. But according to Chalmers, it is precisely this kind of analysis that makes materialism feasible, and if there is a feature of the mind that is not functionally analysable, then physicalism is false.

Even though all these arguments point to slightly different conclusions, they seem to coincide in their logical consequences. Their differences have to do with rhetorical aspects expressed in different examples, and each argument seems to better supported by a particular example. The arguments are supported by certain possibilities not evidently conceivable. Without a further justification, these possibilities show at best that phenomenal consciousness constitutes an epistemic problem, but not a metaphysical one, and the truth of physicalism shouldn't be threatened. After Kripke's *Naming and Necessity*, conceivability cannot be taken as a safe way to possibility. That's why Chalmers develops a semantic framework within which the argument against physicalism is formulated. In order for the previous argument to work, there must be a link between conceivability and metaphysics. We could find out a posteriori that consciousness is identical with a certain brain state, and that it is necessarily so. Let's assume that P is the set of physical facts in the actual world and C is the set of facts about phenomenal consciousness, Chalmers' argument should have the following form:

(1) P & ~C is conceivable.

(2) If P & ~C is conceivable, F & ~C is metaphysically possible.

(3) If $P \ \& \ \sim C$ is metaphysically possible, materialism is false.

(4) Then, materialism is false.

As mentioned before, a clear problem with this argument is that conceivability is not a good guide for possibility. Not everything we conceive is possible. In order to address this issue, Chalmers distinguishes between two types of conceivability: prima facie conceivability and ideal conceivability. Prima facie conceivability is linked to a particular subject; on the other hand, ideal conceivability refers to the conceivability of an ideal subject, free of the cognitive limits of a real subject. Additionally, Chalmers draws a further distinction between positive and negative conceivability: S is negatively conceivable if it is not a priori excluded and S is positively conceivable if a coherent situation can be imagined where S is true. These four categories can be combined, defining four types of conceivability. None of these types of conceivability ensures, however, the reliability of our intuitions regarding what is possible. According to Kripke, there are truths that are necessary and we can only know about them a posteriori. This is the case of an identity statement like "Water is H_2O ".

Chalmers's semantic framework addresses this kind of statements pointing out a double way of assessing the application of our concepts across possible worlds. Each way corresponds to a specific function in which possible worlds are mapped into references. Firstly, possible worlds can be considered as actual and the reference of the concept is fixed depending on how that possible world is. If we were to consider Twin Earth, where the liquid that fills the oceans is XYZ, as a real world, then we would conclude that water is not H_2O , but XYZ. This is what Chalmers calls a primary intension. Secondly, given that water is H_2O in the actual world, we can consider other possible worlds as counterfactual, and assess the reference of our concepts after having fixed it in the actual world. Then, according to this way of assessing the reference of the term "water", water is H_2O in every possible world. This semantic framework is called two-dimensional because the reference of

concepts can be assessed across possible worlds according to two distinct intensions.

In the case of consciousness, it is possible that phenomenal mental states are identical with neural states and maybe this identity can only be discovered a posteriori. Pain could be identical with C neural fibres activity, in the same way as water is H₂O. A conceivability in terms of primary intensions does not entail a conceivability in terms of secondary intensions. But according to Chalmers, secondary intensions are irrelevant in the case of phenomenal consciousness. If we can conceive a possible world where actual physical facts hold but phenomenal consciousness is missing, then physical facts do not exhaust all the facts. An analogous example in the case of water would be to conceive a possible world where actual physical facts hold but there is no liquid that fills the oceans. That's because in the case of phenomenal consciousness, seeming is not different from being. If something feels like a pain, then it is a pain. Thus, in the case of consciousness, primary conceivability allows us to draw metaphysical conclusions.

Instead of addressing the correctness of the two-dimensional semantic framework proposed by Chalmers and the antiphysicalist argument built upon it, I shall locate the discussion in the metaphysical conception of consciousness. The semantic framework just sketched could at best describe the logical consequences of conceiving consciousness according to its primary intensions, but it neither tells anything about the existence of such primary intensions, nor about what consciousness is. The argument relies on the intuition that in the case of phenomenal consciousness "seeming" is "being" and in some sense, regarding consciousness, error is not possible. In the following chapters, the metaphysical conception of consciousness that lies behind Chalmers' arguments will be described and challenged.

Chapter 4 deals with the epiphenomenal position that seems to ensue from Chalmers' concept of phenomenal consciousness. According to the conceivable possibilities regarding consciousness, we could conceive of a zombie world, so that every physical fact would be as in the actual world, but phenomenal consciousness will be missing. However, if this is so, then consciousness does not play any causal

role and cannot be considered explanatory relevant in a clear sense. The causal irrelevance of consciousness amounts to its epiphenomenalism.

In contemporary debates, epiphenomenalism regarding the mind is a way of taking the mind seriously without questioning the validity of physicalism. Part of the content of physicalism is the causal closure of the physical world. On the other hand, the mind seems to be causally autonomous. Functionalism is an attempt to make compatible the causal relevance of the mind and the truth of physicalism. But to what extent is this possible? Let's look at the causal argument in favour of physicalism:

- (1) Mental states have physical effects.
- (2) Every physical effect has as a complete cause a physical state.
- (3) The physical effects of mental states are not overdetermined by distinct causes.
- (4) Mental states are identical to physical states.

The core of the argument is the exclusion of a double causal line, mental and physical, so that mental states and physical states must be identified if the causal efficacy of the mental is to be preserved. The problem is that this manoeuvre is not enough to save mental causation. In addition to being causally efficacious, mental states should cause their effects in virtue of being mental. Functionalism defines a framework within which a specific kind of properties can be studied; each functional property has a certain causal profile. However, functional properties depend crucially on their physical realizations, insofar as they do not possess autonomous causal powers; and that means that only the realizers of the functional properties enjoy genuine causal efficacy. This is what Kim (1998) has dubbed the argument of causal exclusion. But what does it mean to genuinely have causal powers? Mental properties are not the only kind of properties subject to a putative reduction. Biological, chemical and geological properties can be understood as instantiating

functional properties. What can prevent the argument of causal exclusion to expand and call into question paradigmatic cases of causation, and even the very notion of causation? It seems that if the argument of causal exclusion were to generalize, the causal powers of everything would drain away. According to some interpretation of the argument, only the existence of a bottom-level would stop this draining away. On the other hand, if physical properties were the candidates that constitute the bottom level of reality, they should be especial in some sense; they cannot just be dispositional properties playing a causal role.

Even though Chalmers does not address these issues directly, they are relevant to understand the implications of his approach. According to Chalmers, physics can fix the content of physical properties only relationally, but he assumes the basic entities that causation relates are intrinsic, and the only kind of intrinsic properties with which we are acquainted are phenomenal properties. But the dichotomy between intrinsic or categorical properties, on the one hand, and dispositional properties, on the other, is, in turn, problematic. Categorical properties reveal what things are in themselves; an example of this kind of property is the shape of an object or its mass. Dispositional properties express, by contrast, how a determinate object would behave under certain circumstances; fragility is a paradigmatic example of a disposition. This distinction is crucial to understand functionalism and its relation to physicalism, because in the functionalist framework there is a distinction between a functional property and its realizer. Additionally, each realizer can instantiate the proper functional property only if it is put in the relevant context. A way of accounting for dispositional properties is to pin down a categorical property that determines the dispositions a certain object has. Physical properties should be the main candidates for being the categorical basis of dispositions. However, how categorical properties could account for dispositional properties? Firstly, if the relation between categorical and dispositional properties is contingent, then instances of dispositional properties should be individuated independently of the existence of the categorical basis, and the latter ones cannot serve as a criterion of attribution of dispositional properties. Secondly, if there were a necessary a posteriori connection between the categorical and the dispositional property, then

the disposition should be individuated equally independently of its alleged categorical basis. If this is so, why is it necessary to posit the existence of categorical properties? It could be said that the attribution of dispositions has to do with the truth of some counterfactuals: a fragile glass is a glass that breaks if it is thrown. However, when it comes to the attribution of categorical properties it would be difficult to dispense with some counterfactuals. Otherwise, what would it mean that an object has certain mass independently of counterfactuals involving forces and accelerations? Maybe the problem is rather that dispositions should be able to explain their manifestations without being trivial. But in order to meet this requirement we just need a disposition that manifests itself in different ways, so that the disposition is not exhausted by the manifestations it aims to explain.

How can physicalism uphold this analysis? As mentioned, in order for a functional property to be instantiated, a proper realizer must be put in the relevant context. According to physicalism, physical facts fix the rest of the facts of the world, so physicalism requires physical properties to explain why a determinate object instantiates a certain functional property when put in the relevant context. This intuition can be summarised in the twin-substitutivity test: "If A and A* are two physical duplicates, then, for any spatiotemporal context which fixes the possession of a property P by A, had A* been in the place of A, then A* would also have had property P" (Corbí y Prades 2000: 59). But the twin-substitutivity test will not do, according to Corbí y Prades (2000), because it does not fix a set of properties explanatory relevant, given that the test is not linked to manifestations independent of the dispositions it aims to account for. Besides a proper context of realization, functional properties are individuated against some background conditions that we regard as normal. For every conditional that might include in its antecedent a set of physical properties, a context of realization and a set of normal conditions, it wouldn't follow necessarily a determinate output, and then, the idea of a set of basic properties that fix the causal powers of things does not make sense. Causal processes cannot be individuated without any reference to a set of normal conditions fixed by a determinate practice. The metaphysical dichotomy of categorical and

dispositional properties is the result of neglecting the conditions under which we form our concepts and give them a metaphysical status.

Another form of epiphenomenalism arises in Chalmers' account, specifically related to the intrinsicness of phenomenal properties. But in spite of the two different concepts of the mental, Chalmers points to links between phenomenal consciousness and cognition. The clearest example are phenomenal judgements, namely, judgements about our phenomenal experience. Chalmers recognizes that here a paradox arises: on the one hand, phenomenal judgements are about phenomenal experience; on the other hand, they are functional states, and thereby phenomenal consciousness does not play any role when they are uttered. This paradox is the result of accepting the causal closure of the physical world and the irreducibility of consciousness. But then, how knowledge about our own conscious mental states can be justified? The beliefs of our zombie counterpart wouldn't be justified, not even true. But if this is so, to what extent beliefs about our own phenomenal states are justified? Chalmers' answer to this problem is that, in the case of phenomenal consciousness, no causal link is required to have knowledge. The knowledge that relies upon a causal link is fallible, but this does not hold in the case of consciousness. According to Chalmers, sceptical scenarios involving our own phenomenal states are not possible. The relation we hold to our own experience, Chalmers follows, is a relation of acquaintance. This relation does not immediately produce knowledge and it is more basic than belief; even our beliefs about our phenomenal states could be wrong because we don't pay enough attention to them. This introduces a very realistic view of consciousness.

If we take into account the practices according to which we attribute conscious mental states to others, it will be dubious that Chalmers' metaphysics of consciousness is coherent. Chalmers assumes that it makes sense to speak about separate fragments of consciousness. However, it seems that the existence of consciousness in a being entails showing some coherence and stability in behaviour, expressions, actions and bodily manifestations, and the dividing line between what is mental and what is not, is vague. Conscious experience is inextricably linked to

our actions and the metaphysical division epiphenomenalism requires makes no sense.

In Chapter 5, the metaphysical assumptions about consciousness are called into question. It is common in the debates about consciousness that some intuitions are brought up as unavoidable data, immune to objections. However, it will be shown that far from being basic, Chalmers' conception of phenomenal consciousness depends on dubious metaphysical assumptions. As in the case of the categorical/dispositional distinction, there is a way of understanding consciousness that makes sense within a determinate practical framework; but when the concepts that make sense in a particular practical framework are pushed beyond their conditions of application, they do not make sense anymore, and that's what happens with Chalmers' phenomenal concept of consciousness.

The first feature of Chalmers' account of consciousness is its strong realism: even under epistemic ideal conditions, we could be wrong about others' phenomenal mental states. On the other hand, according to Chalmers, we enjoy a special relation to our own phenomenal states that makes our knowledge of them infallible.

An apparently innocuous thesis about consciousness is that the kind of experiences we have is determined by the time and the place where we are located. Conscious experience is conditioned by sensory stimuli, and the sequence composed by a sensory stimulus, a conscious experience and a bodily reaction (an utterance, a movement and so on) seems natural and unquestionable. However, does this principle apply when a very small time-scale is involved? And regarding space, is there a dividing line in the brain that once crossed by a nervous signal separates a pre-experiential path from a post-experiential path? An approach that answers positively to both questions is what Daniel Dennett (1991) calls the Cartesian theatre, understood as the idea that there is a place and a moment where consciousness is given. This idea is linked to materialism, to the extent that it locates consciousness spatiotemporally.

It could be said that Chalmers' position has nothing to do with this position, insofar as phenomenal consciousness is not fixed by a set of physical facts and, in turn, is independent of a particular place or moment. However, Chalmers' position

tries to account for everyday intuitions regarding consciousness and, in this sense, consciousness is embedded in the material world. Chalmers does justice to this idea by acknowledging the existence of a natural relation of supervenience between the physical and phenomenal consciousness, but at the same time, once dualism is embraced, the mind is unavoidably separated from matter. On the other hand, dualism, and especially Chalmers' position, holds a strong relation to time and space. The talk about the "correlates" of experience or the idea that consciousness "arises" in some sense from a physical basis refers explicitly to space and time. The idea of a spatiotemporal locus is not so important, and even empirical evidence is discrediting this position. What is more important is the idea that phenomenal consciousness must have a determinate metaphysical structure if it is to be correlated with the functional aspect of the mind. But then phenomenal consciousness can be fragmented, it has a discrete metaphysical character. Then, realism and discrete character are the two features that implicitly Chalmers assumes regarding consciousness. In the rest of the chapter, I'll show how these metaphysical assumptions are not intelligible in the case of consciousness.

Dennett (1991) analyses the phi phenomenon to show the unintelligibility of the Cartesian theatre. The phi phenomenon is the visual delusion that makes us see motion when indeed there is a very quick succession of images. A delusion based on this phenomenon can be produced when two flashing successive lights are projected in a short period of time. The subjects of this experiment report seeing a unique spot moving from left to right. A variation of the experiment changes the colours of the lights, so the first light is red and the second green. In these variations of the experiment subjects report a change in the colour of the spot in the middle of the trajectory. If we accept the principle that consciousness is linked to a particular place and moment, we have two main options to account for the phenomenon. According to the first one, there is a conscious experience linked to the second light, but the brain, somehow, re-interprets the information it receives and modifies the memory of the past, so that the subject remembers a moving spot changing its colour; this is an Orwellian theory. In the second theory there is not a conscious experience erased from the memory, but there is a delay in the emergence of

conscious experience and an unconscious perception of both lights; later there is a projection of a “composed” experience, elaborated by the brain, where there is a single spot moving and changing its colour, and this is a Stalinesque theory. What evidence could adjudicate between both theories? The problem is that every objective evidence is compatible with both theories. On the other hand, resorting to the first- person point of view will not do, because the description of the experience will be the same in both cases. The difference between both theories is not intelligible, given that we cannot understand which of them is true. That means that conscious experience cannot be linked to a determinate place and moment in every possible situation, as Chalmers argumentation assumes.

Dennett’s position has been objected on the basis that it takes for granted a form of verificationism, but according to my interpretation of Dennett’s argument, what the argument shows is that our commonsense concept of conscious experience cannot be extrapolated to every possible situation and, therefore, cannot be given a metaphysical status for, otherwise, it becomes unintelligible. Dennett’s argumentation cannot be assimilated either to an argument based on a sceptical scenario, to the extent that in the case of a sceptical scenario we cannot know with certainty about the truth of a situation, but we recognise the intelligibility of the two possibilities at stake; in the case of the Orwellian and the Stalinesque theories, we don’t understand the difference between both theories.

Dennett’s argument, based on empirical evidence though, points to a conceptual limit of the Cartesian Theater model. Dennett (2005) presents a similar argument based on a strange disorder called Capgras delusion; in this case, the argument aims to show why qualia cannot make life worth living. The Capgras delusion consists in believing that a loved one has been replaced by an impostor. The belief can lead the person who suffers the syndrome to attack or even kill the person involved in the delusion. One explanation of the disorder refers to an impairment in the emotional recognitional limbic system of the brain: despite having normal perceptive abilities, the sick person has an anomalous emotional reaction towards the subject. Dennett imagines another disorder called the Clapgras delusion, and it consists in suffering an inversion in the affective dimension we

associate with colours, so that Mr Clapgras has his emotions inverted regarding colours (maybe, what he felt when he had a sensation of red is now felt towards sensations of blue and vice versa). If ex hypothesis qualia were to make life worth living, an inversion as the one suggested wouldn't be possible, and the disorder should rather be described as a qualia inversion. However, in the thought experiment devised by Dennett, it seems the proper description of the situation is rather that his colour qualia remained unaltered but these emotions have been inverted. Once more, how could we adjudicate between both positions, given the way consciousness is being conceived in such kind of experiment?

Chapter 6 addresses the question of how it is that consciousness makes life worth living. More specifically, I will argue that Chalmers' conception of consciousness cannot explain why consciousness is important for us and is even incompatible with the idea that consciousness is what matters in our lives.

The chapter starts by analysing three types of unconscious mental states, with the aim of determining what is specific about conscious mental states. Additionally, the study of the unconscious dimension of the mind could shed light on how we relate to our own mental states and what's the proper relation that endows consciousness with the value we attribute to it.

The first case of an unconscious mental state is the disorder known as blindsight. Blindsight is the result of brain injury and the subjects that suffer from it do not see anything in some part of their visual field. However, when they are prompted to guess a figure that had been projected in that part of the visual field, among different options, they answer correctly. It seems that in the case of blindsight, visual information is processed without the existence of conscious experience, at least in the phenomenal sense. Blindsight could be understood as a real example of a zombie. But it is indeed a case where, even from a functional point of view, there is a severe impairment in the behaviour of the subject. On the other hand, the way Chalmer poses the debate closes the possibility of ascertaining whether in cases of blindsight there is phenomenal consciousness, given that in these cases the subject does not report the perception of anything and we cannot

determine whether there are relevant phenomenal facts in spite of the negative testimony of the subject.

Unconscious mental states are not necessarily pathological. There are many functions of living organisms that are beyond their conscious control, even though they are somehow related to the mind. In everyday life there are examples of mentality that are unconscious, although in some cases we cannot tell for sure whether they belong exclusively to the unconscious part of the mind. This is the case of automatic driving, where a subject can drive for miles without remembering having perceived anything. Of course, in this case, there must be some kind of perception, otherwise the driver would have had an accident. But this perception is unconscious in some sense. According to Armstrong (1980), what accounts for the difference between being conscious of our surroundings and merely perceiving things automatically is the existence of an “inner eye” that gives us access to internal states; in the case of unconscious driving, the inner eye would be closed and we can pay more or less attention to our own mental states. We are conscious in virtue of paying attention to an internal state that it is there anyway, independently of our attention. This model of explanation of consciousness is incapable of accounting for what is important regarding consciousness, as I will argue next.

A question that must be answered to clarify these problems is what kind of relation we hold to our own mental states. As an initial approach, we enjoy a special authority regarding our own mental states, so we are the best persons to ask when it comes to know about our own minds. What could it be our relation to what Chalmers calls phenomenal states? And why phenomenology makes life worth living? According to Chalmers, this relation is acquaintance, and it is characterised for being a kind of non-causal direct relation to our mind. In order to show the limits of Chalmers’ conception of consciousness I’ll talk about a third kind of unconscious mental states. These are the states that, according to the psychoanalytic theory, are linked to some kind of repression and disrupt the agent’s behaviour. Regarding the phenomenal concept of the mind, it seems there cannot be these kind of states, because I cannot not be unaware of my phenomenal mental states. But how can this authority be accounted for?

Finkelstein (2003) defines an account of the authority of the first person point of view that is called detectivism. According to detectivism, I am conscious of my mental states thanks to a cognitive process that allows me to detect their presence. Finkelstein distinguishes, in turn, between two forms of detectivism. The first one, old detectivism, posits an inner sense that only metaphorically can be understood as working like the external senses. The relation to our own mental states is non-inferential and we know about them with certainty. The immaterial character of this kind of inner sense and its consequences would be the strongest objection against old detectivism. The second form of detectivism, called by Finkelstein new detectivism, posits a cognitive process of detection that is on all fours with the external sense, so there are no essential differences between them. But would this new form of detectivism account for the authority we enjoy regarding our mind? In principle, empirical research should address this issue at some point, so this is not a strong objection against new detectivism. It seems that detectivism can always be restated to block objections.

Finkelstein poses an objection to detectivism that attacks its core tenet, in all its versions. As we have seen, some of our unconscious mental states were linked to a personal mental conflict that represses a certain mental state, although it continuously manifests itself in anomalous ways. An example of this is being unconsciously angry with somebody. A detectivist position posits a detection mechanism that allows us to know about our mental states. But according to Finkelstein one thing is to know one's mental states and another thing is having them consciously. I could be "conscious of" my anger, without being "consciously" angry. To know about somebody else's anger, we need evidence, but we don't relate to our own mental states by means of evidence. The problem is that detectivism cannot account for that difference because it relies upon the idea of detection. Resorting to the specificity of phenomenology or to the particular way we perceive things through senses will not do, either, because the mere detection does not allow us to deepen the understanding of the authority of the first person. Additionally, detectivism cannot help us to understand what matters about consciousness.

The objection to detectivism affects Chalmers' account of consciousness, as well, to the extent that the latter also points to a discrete item which by itself should account for the authority of the first person. For Chalmers, consciousness is the instantiation of phenomenal properties, which have an intrinsic character. But this intrinsic character does not help us to understand why consciousness is valuable.

Chalmers thinks about phenomenology as if phenomenal states were to have a rich set of properties which are there anyway. This is what Alva Nöe (2004) calls the snapshot conception of experience. According to this conception, our experiences are objects with well-defined properties. But what empirical evidence shows is that we are not so good at describing our experiences, even though we believe there is a complex set of properties in our experiences which are there anyway. It seems there is a delusion regarding what we take our experience to be, but, according to Nöe, the richness of phenomenology is not but the richness of aspects of the world. Then, if this way of thinking about phenomenology is correct, our experience cannot be conceptually independent of our body and of the way it relates to its surrounding. What makes a difference in the several ways of perceiving the world is not a difference in intrinsic phenomenal properties, but rather the distinctive way in which our body can relate to the world. The idea of an intrinsic source of value is unintelligible, insofar as the idea of value is linked to the existence of a being embedded and engaged in the world in a particular way.

The rest of the chapter deals with Chalmers' science of consciousness. Even though Chalmers thinks that physicalism is false and physical facts do not fix the facts about consciousness, he thinks the ontological catalogue of science could be extended to include consciousness among natural phenomena, and a science of consciousness would be feasible. The basic principle of this science would be the coherence between phenomenology and cognition. According to this basic principle, our judgements about our own experiences are reliable, and there is a correlation between phenomenology and cognition or psychology. Additionally, phenomenal consciousness has certain structure that corresponds directly to the structure of awareness (the functional concept of mind). This science takes for granted some of the assumptions that have been called into question in the previous chapters, like

the idea of correlation between phenomenology and the possibility of individuating discrete fragments of experience.

Even though Chalmers presents his position as an extension of natural science, it is dubious that he is successful in his attempt. The hard problem of consciousness is *sui generis* and the appeal to plausibility in this case is no use, insofar as, in case we accept the terms of the debate posed by Chalmers, we even don't know what would count as a solution to it, so the standard notion of plausibility cannot be used here.

Finally, could there be a science of consciousness that give an answer to the problems linked to it? The problem of consciousness is stated within the framework of the search of an absolute conception of reality, even though Chalmers' position acknowledges this project is a blunder. His attempt to conceive of a unified science collides, precisely, with the absolute conception of reality. The unity of science can be understood either as a kind of internal connection between the different sciences, not necessarily reduction, or as a methodological unit. Chalmers' position oscillates between the two senses of unity: Chalmers' antireductionist arguments seem to point to the difficulty of finding an internal connection between the different sciences, but this is an independent result from the impossibility of having a unified science from a methodological point of view. What I've tried to show is that consciousness is not a "thing" that can be studied among other things, but it depends on the different ways beings with a particular body relate to the world, so it is difficult to understand how a science could come up with everything worth saying about consciousness.

Introduction (Introducción)

The problem of consciousness is considered to be one of the most baffling puzzles and seems to constitute a serious obstacle to a mainstream theoretical project¹ which has shaped to a great extent the philosophical debates of the last decades. It is a problem that can be posed by means of a few simple grabbing images which do not resort to any specialized knowledge, but for some philosophers no future scientific development will be able to cope with it. In a few words, the problem of consciousness can be described as the difficulty to understand how neural processes can give rise to conscious experience. This formulation has a clear epistemological character, but it can be rephrased metaphysically by saying that even if a clear set of facts that are indefectibly associated with consciousness obtained, conscious experience could still be missing.

The easiness with which the problem of consciousness can be formulated is linked with its degree of generality; conscious experience is a ubiquitous mental phenomenon that encompasses different kinds of states: sensations, feelings, emotions, thoughts, beliefs, intentions and so on. Of course, a great deal of the debate about consciousness has to do with the issue of clarifying the different concepts of consciousness, and what was meant to be a rather general problem transformed into a series of different questions, not all of them clearly philosophical. This ambiguity is accompanied by a manifold of answers aiming to cope successfully with the problem of consciousness. However, according to some philosophers, those theories have not even taken a step in the direction of addressing it properly and have misfired the real target.

This dissertation deals with David Chalmers' approach to the problem of consciousness. Chalmers' theory of consciousness has been most influential, for three reasons at least. Firstly, Chalmers distinguishes between different kinds of problems related to consciousness and identifies a proper philosophical question that is not to be conflated with the different issues associated with the research

¹ To some philosophers, like Jaegwon Kim, the problem of consciousness is the greatest difficulty for the physicalist project.

program of cognitive science. Secondly, his approach is couched in a language that allows locating it in the geography of different contemporary debates and, as a consequence of this, the relation between the debate on consciousness and other discussions can be assessed. Finally, Chalmers' formulation of the problem of consciousness is amenable to rational discussion, to the extent that it is fixed in a set of explicit statements.

This dissertation aims at analyzing the assumptions on which Chalmers' approach to consciousness relies in order to challenge them. Additionally, as a result of this criticism, the terms of the debate will be defined and the standard geography of these issues will be rejected. Current philosophical debates tend to develop according to three general stages. In the first place, a new problem is posed (sometimes an old problem is couched in a new terminology) and a few accounts try to solve it; next, some general solutions to the problem develop, shaping a general paradigm of discussion that will be taken seriously by the academic community; finally, this paradigm become complex, entering a manneristic phase where the relevant literature deals with very specific issues and each of the accounts ends up being a real field of expertise. In this latter phase, it cannot be said there is a clear winner approach, at least for a long time, and the discussion stagnates and languishes. At the same time, inside each paradigm there can be several sub-approaches that, in turn, can be divided in different "isms".

Some examples of these general paradigms are the compatibilist-incompatibilist approaches to free will, the causal *versus* the descriptivist theories of reference or the foundationalism vs. coherentism debate in epistemology. There are, in turn, several kinds of compatibilists, realists or foundationalists and there are no limits to the different subcategories that can be coined to map these philosophical discussions. The debates on the philosophy of mind, and especially on consciousness, are not an exception and for each discussion a host of well-entrenched accounts are to be found. During the last decades, the discussion about conscious experience has become increasingly complex and very sophisticated conceptual tools have been developed in order to cope with the different difficulties plaguing the debate. All this technical apparatus is supposed to account for the

apparently simple intuitions on which all the discussion relies; however, it is not clear at all that the new concepts are less obscure than the initial ones. Constructing very complex conceptual buildings and drawing conclusions out of them does not suffice to provide a true theory (as philosophy in Middle Ages clearly shows). Eliminativism, functionalism, identity theories, representational theories, dualism, these are some of the most prominent philosophical approaches to mind, and as I will show, some philosophical problems cannot exist without a general framework shared by those theories.

Another consequence of this way of doing philosophy, and not less important, has to do with how the largest part of the related literature deals with highly abstract discussions, remotely connected with the original perplexities that triggered the debate. Even in case the discussion came to an end and one of the approaches were to become victorious (if this is possible at all in philosophy), what would put an end to the debate? It seems that only a very subtle argument, maybe far away from the original context that generated the original perplexity, could be a candidate capable of adjudicating between the different accounts. In this sense, it could be possible that the initial perplexity remained, and only a very hard-to-follow-chain of reasoning could lead us to think that the answer to the problem of consciousness is one instead of another. But is this the purpose of philosophical reflection? Is the original impulse to philosophize satisfied by the possible result of the activity just described? Of course, this is not to say that philosophical reflection cannot take us far away from a starting point; but the continuity of a philosophical journey should be recognized as illuminating, so that the initial perplexity is not haunting us anymore or, at least, there is a minimal understanding of this perplexity. Surely, sticking to a false position is not the only kind of failure in philosophy.

A well-known example is the Augustinian view of language, according to which understanding how language works amounts to understanding that words stand for things. Additionally, knowing the meaning of a word is knowing the thing that word stands for; sentences are the combination of these words, which are names for objects. The reason why this position is not philosophically satisfactory is not its falsity, but rather its limitations when it comes to understanding the way

language works. Of course, we could explain language by saying that each word stands for something, and in some sense this is true. The word “table”, when properly used, stands for a table; and the word “belief” normally stands for a belief, even though it is not clear in which sense a belief is a thing in the same way a table is. What about the word “between”? Could we say that “between” refers to something? Or is it rather the way two things relate to each other? But if this is the case, is it a mistake to say that a relation is a thing we are referring to? Another problematic example could be an adverb, like “quickly”. An adverb refers to a way of doing things, and in this sense, we could engage in the discussion on whether a way of doing things is, in itself, a thing or not. But which kind of data could help us to adjudicate between both positions? Can there be a solution for this problem which is not a matter of stipulation? The problem with the Augustinian conception of language is not its falsity; the problem is, rather, that it does not help us to understand how language works. This explanation cannot tell us much about what we are doing differently when we say “I promise not to deceive you” and “The Earth is not flat”. According to such a conception, we are just referring to certain items that relate to each other in a certain way, but we cannot say this explanation is false.

It is not clear to what extent a philosophical enquiry should consist in gathering an exhaustive set of interrelated propositions regarding any true thing (or at least, any true basic thing) to be said about a subject, so that any relevant fact related to that very issue can be subsumed in this set, or else be a logical consequence of it. If this were the only purpose and end of philosophical reflection, the history of philosophy would be the history of a failure and it wouldn't be much interesting to read classic philosophers, insofar as none of their theories can be considered as true. In this sense, the interest in reading contemporary philosophy should be on all fours with the interest a scientist might have in reading the more recent papers about her field of expertise, and the history of philosophy would relate to philosophy in the same way the history of science relates to science. At least it is clear that in the practice of science old works are rarely important, whereas in the case of philosophy a new approach to a classic work can be much relevant in contemporary discussions.

But then, what else makes philosophy worth studying? As said before, philosophy should be illuminating, and in order for something to be illuminating the interests of human beings should be taken into account and some kind of fulfillment regarding human expectations should be attained. These remarks are to be held for every human activity, even though they are often neglected. During centuries, the ideal way of performing a theoretical activity had to do with being detached from the object of study. The purpose of this antiseptic recommendation was to avoid any kind of bias which hailed from human subjectivity; however, when this practice is understood on the assumption that an object in itself can be known, without any single influence of human beings, we are going beyond a methodological practice, entering the zone of dubious metaphysics. This project of attaining knowledge of things in themselves, removing any trace of human subjectivity, is what Bernard Williams called the absolute conception of reality.² Regarding consciousness, these remarks are especially relevant, given that consciousness is the subjective phenomenon *par excellence*. But how can something essentially subjective be understood as being completely objective? The absolute conception of reality owes its plausibility to the astonishing success of natural science during the last centuries and to the explanatory power of this science. It is a project based on a success story, maybe without precedents. More especially, the impact of Newtonian mechanics on modern philosophy can be hardly overstated. Newtonian mechanics became the ideal model of human knowledge and the most important works of modern philosophy present themselves as laying the foundations of knowledge, or else point to some basic elements to which everything that is true reduces and try to parallel the achievements of modern physics when it comes to understanding human beings. It goes without saying that no similar success has been reached in further areas of human knowledge. The mind can be conceived of a set of separate atoms (ideas, mental states) that behave according to certain strict laws, like bodies subject to the laws of physics. This way of thinking about the mind can work as a framework to many research projects and can be even fruitful to some extent, but it is unlikely this is the whole story to tell.

² See Williams (1978).

But then, what makes a philosophical approach satisfactory in addition to its truth? Maybe it cannot be said much about why a philosophical approach is illuminating independently of the particularities of a concrete discussion. The approach of this dissertation does not put an emphasis on the truth or falsity of Chalmers' position. Rather, the arguments presented in this dissertation aim at achieving three kinds of results which not necessarily collide with the truth of Chalmers' account: firstly, I intend to make clear the general assumptions that make possible the formulation of the problem of consciousness; secondly, taking into account these assumptions, I will argue that Chalmers' cannot draw intelligibly the conclusion he is supposed to reach; thirdly, even though Chalmers' theory intends to say everything philosophically interesting regarding the problem of consciousness, it will be called into question that this is the case, sketching core aspects of the understanding of consciousness that lie beyond the scope of his theory. Even without a clear conception of what an illuminating philosophical approach is, I assume that a philosophical enquiry delivering these three results, or some of them, could be illuminating enough. The degree to which this objective is accomplished, meeting the theoretical expectations generated by the puzzle of consciousness, is to be assessed in the particular development of the discussion. At some point, if an approach does not seem to be illuminating to someone after having been properly understood, then not much more can be said in its favor.

Chapter 1 introduces the physicalist framework within which the problem of consciousness is posed. To begin with, the determination of the content of physicalism in terms of physics (either current or future) is of no use. Chalmers does not get into detail either when it comes to determine the content of physical properties. The emphasis of Chalmers' definition of physicalism is put in the logical hierarchy of reality, where physical facts (whatever they are) are the basic facts of reality. Chalmers' account is committed to a layered conception of reality, where physical facts are regarded as the basic facts of the world insofar they are located in the fundamental level of reality, so that once the physical facts of the world are determined the facts of the world are determined *simpliciter*. For Chalmers, however, consciousness cannot be determined with logical necessity out of the totality of

physical facts. According to Chalmers, every explanation must stop at some point, and some brute facts holding an arbitrary relationship between them will be found at the bottom level. Then, giving an illuminating explanation of something means just explaining how the seemingly arbitrary relations holding between some set of facts are not different from the arbitrary relations linking the brute facts of the bottom level. But even in the case of consciousness we would lack this kind of “illumination”. Physicalism overlaps partially with the absolute conception of reality. At least, it shares with the latter the attempt to find a privileged and fundamental theoretical perspective that stands in a special and asymmetrical relation to every other perspective. Phenomenal consciousness, according to Chalmers, holds a relationship to the rest of reality that cannot be explained by means of other scientific connections. At the same time, the problem of consciousness is unavoidably formulated within a mainstream physicalist framework. But mind and, especially, consciousness seem to be recalcitrant to physicalist reduction. If physicalism is to be a satisfactory account of reality, it should be able to accommodate the different perspectives about reality, or else reject them as false beliefs or delusions. The latter presupposes what an unmasking explanation is, in terms due to Barry Stroud (2000). Unmasking explanations are common in the history of modern science. There is a basic constraint this kind of explanation must satisfy: the *explanandum* must be identified without presupposing its truth. This is a very demanding challenge for physicalism, which must account for mentality out of the austere terminology of physics. Additionally, the mind is not simply correlated to the world; there are intelligible links between mind and world that go beyond a mere brute connection. A project that rather aims to be illuminating and to build cognitive unities is akin to Bernard Williams’ notion of “making sense”, according to which there are different kinds of narratives that make sense to distinct audiences with disparate interests. Chalmers’ approach cannot be illuminating because it intends to explain consciousness as if such kind of understanding could be reached without any reference to particular audiences or interests.

Chapter 2 analyzes the distinction between the two concepts of the mental that go together with the two kinds of problems about consciousness, namely, the

easy problems of consciousness and the hard problem of consciousness. The psychological concept of mind has to do with the causal explanation of behavior; on the other hand, the phenomenal concept of mind relates to the experiential character of consciousness. Phenomenal consciousness is the one that poses a proper philosophical problem (the hard problem of consciousness). The problem can be stated as the lack of conceptual links between phenomenal properties and physical properties (or properties amenable to be reduced to physical properties by means of a functionalization). I will argue that this problem cannot be stated coherently. Chalmers accepts that the link between psychological (functional) and phenomenal properties is empirical, but in order to establish this point, the content of phenomenal properties must be individuated independently of any functional property, and in this sense there are two possibilities at hand. Either instantiations of phenomenal properties are pinned down resorting to the psychological aspect of the mental, which is co-occurrent with the phenomenal aspect, or are individuated privately, from the first person point of view. Resorting to the psychological aspect of the mental means that a functional property, allegedly co-occurring with the proper phenomenal property, could be chosen in order to determine that a phenomenal property is being instantiated. But the problem with this answer is that there is no clear criterion that guarantees the instantiation of such functional property whenever the phenomenal property obtains. The path of the first person point of view is not more promising. In the rest of the chapter, I examine a Peter F. Strawson's argument to this purpose (see Strawson 1959). In order to speak about the relation I hold to *my own experiences* I should have been previously able to determine what the subset of *my experiences* is, so that they can be distinguished from other people's experiences, but how could I do this taking my private experience as a starting point? How could I point out to the subset of my own experiences without being previously able to ascribe experiences to other people? If the acquaintance with one's experiences has logical priority in order to ascribe experiences from a third person point of view, as Chalmers assumes, then the content of these properties cannot be fixed, and the distinction between the two concepts of mind cannot be drawn coherently.

Chapter 3 deals with the arguments against the reducibility of consciousness. These arguments can be sorted depending on the kind of conclusion they lead to, and on the way their premises may support the conclusion. The first kind of argument is characterized by the drawing of a metaphysical conclusion out of the conceivability of a situation. The most important argument of this kind is the conceivability of zombies, according to which it is possible to conceive beings identical to us in almost every respect -except regarding phenomenal properties- that, nonetheless, lack consciousness. Then, if this is possible -the argument continues- materialism is false. According to the experiment of the inverted spectrum, the second argument of this type, it is metaphysically possible that two identical individuals in every non-phenomenal aspect have their color-experiences inverted. This argument states the falsity of materialism, even though it is not committed to the possibility of the zombies. I will argue the lure of this argument has mainly to do with some rhetorical aspects specifically associated with the case of colors; when the argument is formulated with regard to other kinds of properties, it sounds much weaker. Jackson's Knowledge Argument resorts to the thought experiment of Mary, the scientist that knows every fact about color but never had color perceptions. Once more, it seemed that the argument relied heavily on the intuitions linked to the chromatic experience, but it does not work so well when it is couched in terms of other phenomenal properties, so that its scope is remarkably restricted. Finally, the last kind of argument has to do with the analysis of mental concepts, and the absence of a functional analysis in the case of phenomenal properties. This argument assumes that the absence of a functional analysis, in the case of conscious mental states, is an absence of analysis *tout court*; on the other hand, it is unclear why from the absence of analysis a conclusion regarding the falsity of materialism in terms of supervenience should follow.

Chalmers develops a general argument that synthetizes the intuitions of the arguments presented. The basic idea behind this argument is that we can conceive of a situation where the actual physical facts hold, but phenomenal consciousness is missing. If this possibility is conceivable, the argument follows, then materialism is false. In order for this argument to work, something needs to be said about the

relationship between conceivability and metaphysics, and for this purpose, Chalmers develops a two-dimensional framework that gives support to his argument. Following the two-dimensional theory of meaning developed by Chalmers, there are two ways of assessing the application of a concept through possible worlds: firstly, the relevant possible worlds can be considered as real, and then it is required a previous conception of that thing whose identity is to be determined. This is what Chalmers considers a primary intension. Secondly, alternative possible worlds can be considered as counterfactual, so that the reference of the term is fixed in the real world. According to Chalmers, in the case of phenomenal properties only primary intensions are relevant to assess possibilities. Hence, what is conceivable regarding phenomenal consciousness is relevant for the determination of metaphysical possibilities and a counterfactual situation involving phenomenal consciousness is metaphysically possible. However, the two-dimensional framework is not but a tool that allow us to describe the two relevant intensions in case they existed. On the other hand, the argument relies ultimately on the idea that, in the case of phenomenal consciousness, “seeming” is “being”. In Part II, these further assumptions regarding the especial status of phenomenal properties will be challenged.

In Chapter 4, I explore the epiphenomenalism that seems to ensue from Chalmers’ theory. The irreducibility of phenomenal consciousness in conjunction with the reducibility of every further mental aspect leads to the conclusion that consciousness is causally irrelevant. As discussed in Chapter 1, Chalmers assumes that there can be a mapping between functional properties and phenomenal properties, given their co-occurrence. Even though Chalmers does not call into question the causal efficacy of functional properties, functionalism has to cope with a general difficulty. Functionalism is attractive because it is capable of explaining how mental properties relate to physical properties, precisely because physical properties realize mental properties. But that means that the causal powers of a mental property are the causal powers of its realizer. If we follow this line of reasoning, it seems that only physical properties have genuine causal powers. But the problem does not stop here: it is not clear whether physical properties indeed

capture causal powers. What is special about physical properties that would enable causation? It seems that, unless there were a bottom level of reality, the causal powers of things would drain away. Chalmers does not address these questions directly, but he is committed to the existence of intrinsic properties as a fundamental substratum of reality and phenomenal properties are the only ones of which we know about their intrinsicness. It seems that resorting to intrinsic properties solves many metaphysical problems. Physicalist functionalism seems to lead necessarily to the idea that physical properties should be intrinsic in some sense, if they are to be special in the sense other properties are not. I'll try to show that intrinsic properties are devoid of explanatory power and the distinction between intrinsic and dispositional properties only makes sense within a context of normality; consequently, the putative intrinsic character of physical properties becomes unintelligible.

Another form of epiphenomenalism has to do with the specificity of phenomenal consciousness, insofar as the lack of phenomenal consciousness wouldn't make a difference detectable by us in a zombie case, not even when zombies talk about their phenomenal consciousness (this is what Chalmers calls the paradox of phenomenal judgement). According to Chalmers, the content of phenomenal properties cannot be individuated by their causal role, so that the epistemic relation we hold to their instantiations must be different; this is the relation of acquaintance. This relation is not positively characterized, but is rather an *ad hoc* movement, forced by the dialectic of the discussion. Negatively, acquaintance can be defined as an epistemic relation, such that no sceptical scenario is even intelligible regarding consciousness. Chalmers' metaphysical conception of consciousness assumes that it makes sense to talk about separated fragments of consciousness. I will argue that part of understanding that a being is conscious has to do with how mental states are inextricably linked to complex and stable ways of acting, so that conscious experience cannot be segmented in isolated fragments.

In Chapter 5 I examine, specifically, the discrete and realistic conception of phenomenal states Chalmers is committed to and, once again, I challenge its intelligibility. These principles involve concepts that are innocuous in everyday

practices, but fail to make sense when applied out of the normal framework, where they become metaphysical. One of these principles is the idea that there is a spatial place (in the brain) where consciousness occurs. This is what Dennett calls the Cartesian Theater. Even though, in principle, almost nobody currently subscribes to the idea of a *locus* where consciousness takes place, the idea of a correlation between psychology and phenomenology assumes that it should be possible to link tightly phenomenal properties to functional properties (on the assumption that the latter are amenable to be located in space). But if this kind of correlation is possible, instantiations of phenomenal properties could be isolated in a discrete way. The second metaphysical principle assumed by Chalmers is that for every question regarding the instantiation of a phenomenal property there is a definite answer; this is a strong realistic conception of consciousness. I will support Dennett's argument in *Consciousness Explained*, according to which neither the conception of the Cartesian Theater or the realistic assumption regarding instantiations of phenomenal properties make sense. Dennett's argumentation is not to be conflated with a sceptical or verificationist argument. However, the conclusions of this chapter leave some philosophical dissatisfaction, insofar as no alternative approach is presented. This dissatisfaction is partially compensated for in Chapter 6.

In Chapter 6, I describe the limits Chalmers' account has regarding its conceptual power to explain the importance consciousness has to human beings. Chalmers' theory of consciousness falls short of explaining why consciousness matters to us and why allegedly it makes life worth living. The hard problem of consciousness is elusive and for any attempt to solve it can easily be rephrased in terms of a further aspect of experience that cannot be reduced to some other kind of thing or property. Instead of providing a standard answer, subject to the same kind of reply, the strategy I will follow seeks to highlight the explanatory limits of the conception involved in the formulation of the problem. Accepting that consciousness constitutes a set of facts devoid of any conceptual link to the rest of reality has the undesirable consequence that there is no way we can understand the valuable role consciousness plays in our lives. Another consequence of this approach is the

impossibility of determining the relevant facts that would allow us to settle the question.

Chalmers assumes that a conscious being holds a relation of acquaintance with its own mental states, and that's what makes them conscious. This relationship of acquaintance cannot explain why consciousness is valuable for us, insofar as it functions as a *magical* relation. These difficulties show the limits not only of Chalmers' approach, but of any theory that interprets consciousness as the instantiation of some kind of particular properties and seeks to grasp these properties as isolated items of study. Even though it is difficult to deny that when someone is conscious of something, some property is detected in a certain way, this movement by itself fails to illuminate our understanding of consciousness. A better framework to understand consciousness should take into account the multiple relation between body, mind and environment and instead of facing a mystery, the philosophical study of consciousness should take into account how these factors interrelate in different constellations. On the other hand, phenomenology is not to be understood as something detached from the outer world, constituted by a variety of properties that hold a contingent relationship with the world. Rather, I suggest the alleged richness of phenomenology is not but the richness of the world. For Chalmers, being conscious amounts to the instantiation of some set of phenomenal properties, but this model of explanation cannot account for the value we attribute to consciousness and it does not say anything about the way consciousness relates to other fields of our life. An alternative account of consciousness is proposed that tries to tie consciousness to our body and to the way our body relates to the world.

The limits of Chalmers' account are reflected too in his sketch of a science of consciousness, based on the coherence between psychology and phenomenology. The explanatory power of Chalmers' would-be science of consciousness is criticized and some remarks are made on the minimal explanatory demands a science of consciousness should meet. Phenomenal properties seem to be connected with the rest of the world in a brute way and this is not a bruteness that could be discovered by means of empirical research; thought experiments allegedly show the especial way consciousness relates to the world. In this sense, the notion of plausibility

Chalmers appeals to is no use, insofar as the science of consciousness he is proposing cannot be understood by analogy with other sciences. On the one hand, it is unknown what would count as an answer to the hard problem of consciousness but, on the other hand, the notion of plausibility Chalmers uses is clearly imported from the context of natural sciences. The hard problem of consciousness is posed within the framework of physicalism. However, Chalmers proposal can hardly be considered naturalistic, insofar as it addresses the question of the methodological unity of science and shows, at best, that the methodological standards of natural sciences can be applied, in some sense, to the study of consciousness. On the other hand, Chalmers is anti-reductionist regarding consciousness and, in this sense, it is unclear to what extent there is continuity between natural sciences and the science Chalmers proposes. If a science of consciousness is to tell us something relevant, consciousness should not be understood as a single phenomenon, amenable to be studied by a single discipline, but as a complex set of interactions between body, mind and world that requires different points of view to be properly understood.

1. Un marco general

En este capítulo, analizaré los rasgos del marco general en el que Chalmers situará la discusión del problema de la conciencia y describiré la concepción metafísica que subyace al modo en que lo formula. El punto de partida de Chalmers es el intento de alcanzar una concepción fisicista de la realidad, si bien la experiencia consciente se presentará como un obstáculo insalvable para completar dicho proyecto, de manera que, planteado en estos términos, el problema de la conciencia se encuentra inevitablemente ligado al fisicismo. Si la argumentación de Chalmers es válida, entre los hechos físicos y la experiencia consciente hay un hueco ontológico que implica la falsedad del materialismo.

Chalmers articulará el fisicismo en torno a la relación de supervinencia y, más concretamente, de supervinencia lógica, de modo que, si el fisicismo es verdadero, los hechos físicos fijan, de un modo necesario lógicamente, el resto de hechos del mundo. Ahora bien, aun si el fisicismo fuera verdadero, todavía habría que responder algunas cuestiones relacionadas con los aspectos explicativos del mismo. Chalmers reconocerá que el modelo de explicación reductiva que propone no es iluminador, lo que significa que el modo en que Chalmers concibe el fisicismo es insatisfactorio en la medida en que no nos ayuda a comprender cómo la experiencia consciente se relaciona con el resto de la realidad. Es evidente que la plausibilidad del fisicismo deriva en gran parte del poder explicativo de la ciencia moderna y es, precisamente, el éxito de la ciencia moderna lo que ha alimentado la aspiración de que puede llegar a darse con una descripción privilegiada de la realidad que nos dice cómo las cosas son con independencia de nosotros. La concepción absoluta de la realidad de Bernard Williams resume el intento de dar con tal descripción y, en este sentido, coincidirá con las pretensiones del fisicismo. Sin embargo, como mostraré, el proyecto del fisicismo, por su propio carácter, no puede dar cuenta de un modo satisfactorio ni de la realidad mental, ni de la importancia que ésta tiene para nosotros, siendo este último elemento el que más propiamente se asocia a la idea de que una explicación es iluminadora.

1.1. Físicismo y supervinencia

El debate básico que permea las discusiones en filosofía de la mente y que actualmente les da forma se articula en torno al físicismo³ o materialismo⁴. El físicismo es la tesis que afirma que todo es físico. Ahora bien, esta definición inicial no atrapa un contenido filosóficamente interesante y, así expresada, es una tesis muy poco plausible. Por una parte, el alcance de “todo” debe ser delimitado de algún modo, ya que hay muchas cosas de las que no podemos decir, de un modo intuitivo, que sean físicas. Por ejemplo, ¿son físicos los números naturales, la ONU, una ley agraria o el valor estético de un cuadro? Esta cuestión es reconocida en la bibliografía al respecto como la cuestión de la completud.⁵ Por otra parte, tampoco está muy claro a qué nos referimos con “físico”. El célebre dilema de Hempel parece situar la discusión en un callejón sin salida. Según el dilema, la referencia al vocabulario de la física actual no nos será de gran ayuda en la medida en que la física actual no es completa. Tampoco podemos apelar al vocabulario de una hipotética física completa del futuro por dos razones: en primer lugar, puesto que no disponemos todavía de dicha teoría, la tesis físicista carecería de un significado determinado; en segundo lugar, aunque asumiéramos la idea de que los futuros cambios de la física mantendrán una continuidad con la física actual respecto a ciertos rasgos generales, no podemos identificar dichos rasgos generales puesto que no tenemos un conocimiento de la física futura y, por tanto, la definición de físicismo se convertiría en trivialmente verdadera.⁶

³ En Stoljar (2010) se encuentra una defensa de esta idea, según la cual la mayoría de las discusiones en filosofía constituyen *placement problems*. A grandes rasgos, lo que Stoljar defiende es que la tesis físicista constituye una cosmovisión a la que no estamos dispuestos a renunciar, pero que entra en conflicto con dimensiones importantes de nuestra vida (el estatuto de lo mental, la moral, la modalidad). El físicismo ha llegado a convertirse en un lugar común en la medida en que ha dado lugar a una forma de hacer filosofía consistente en repensar aquellas dimensiones de la realidad que colisionan con la cosmovisión físicista con el fin de mostrar que, en última instancia, tal incompatibilidad es aparente. Más adelante defenderé una posición semejante, aunque consideraré que el proyecto del que habla Stoljar tiene unos rasgos más determinados que los que él reconoce.

⁴ Si bien ambos términos no coinciden históricamente, en las discusiones actuales son utilizados como términos equivalentes y, a lo largo de este trabajo, los usaré indistintamente. Del mismo modo, utilizaré indistintamente los adjetivos físicista y materialista.

⁵Véase Stoljar (2010).

⁶ Este es un punto señalado por Stroud (2000).

La discusión en torno a la definición positiva y precisa de fisicismo es compleja y no podemos tratarla con gran detalle, pero al menos podemos considerar las posiciones que son claramente incompatibles con el fisicismo con el fin de determinar cuáles son los rasgos más básicos e irrenunciables de toda posición fisicista, en qué medida la experiencia consciente supone una dificultad para el fisicismo, y cuál sería la forma general que una solución al problema de la conciencia debería tener para un materialista. Dejaré de lado la cuestión de determinar a qué nos referimos con físico, pues resultará irrelevante para lo que resta de discusión. Podemos apelar intuitivamente al tipo de propiedades que reconoce la física actual y al tipo de inteligibilidad que ésta nos proporciona respecto a los fenómenos naturales. Sólo reconocer que hay un sentido en el que este tipo de comprensión puede extenderse a todo lo demás constituiría un avance notable. Además, una reformulación más plausible de fisicismo que tenga en cuenta este objetivo podría dar contenido a la idea de que los hechos físicos son básicos, de modo que todos los hechos del mundo dependen de los hechos físicos, en un sentido en el que los hechos físicos no dependen de otro tipo de hechos. Sin embargo, esta afirmación es todavía demasiado general si ha de servir para definir de un modo sustantivo una posición fisicista. Una posición dualista es, en principio, el ejemplo paradigmático del tipo de teoría que sería incompatible con el fisicismo; pero un dualista no tiene por qué comprometerse con la negación de una relación de dependencia de lo mental respecto a lo físico. Hay numerosas relaciones de dependencia. Un tipo de dependencia legaliforme, por ejemplo, que vincule eventos físicos con eventos mentales es perfectamente compatible con un dualismo de propiedades, de manera que es necesario dar con una relación que sintetice lo que es relevante en una posición fisicista. El tipo de relación fundamental desde la que Chalmers definirá el marco general en el que se unificará y discutirá el problema de la conciencia es la relación de supervinencia:

La noción de supervinencia formaliza la idea intuitiva de que un conjunto de hechos puede determinar completamente otro conjunto de hechos. Los hechos físicos del mundo parecen determinar los hechos biológicos, por ejemplo, en la medida en que

una vez todos los hechos físicos del mundo estén fijados, no hay espacio para que los hechos biológicos varíen. (Chalmers 1996: 32)

El esquema básico del que se servirá Chalmers para definir la relación de supervinencia es el siguiente:

Las propiedades B supervienen sobre las propiedades A si no hay dos posibles situaciones indiscernibles respecto a sus hechos A que sean diferentes respecto a sus hechos B. (Chalmers 1996: 33)

Ahora bien, dependiendo de cómo se complete este esquema, pueden definirse distintas nociones de supervinencia, tanto en lo que se refiere a qué entendemos por “situación” (podemos hablar de individuos o mundos) como por el tipo de posibilidad (lógica, empírica). Si al hablar de situaciones nos referimos a individuos, se tratará de una relación de supervinencia local, de manera que las propiedades B supervienen sobre las propiedades A si dos individuos idénticos respecto a sus propiedades de tipo A son indiscernibles respecto a sus propiedades de tipo B. Por otra parte, si hablamos de mundos enteros y de sus propiedades, diremos que las propiedades B supervienen globalmente sobre las propiedades A si dos mundos indiscernibles respecto a sus propiedades A lo son respecto a sus propiedades B.

En lo que se refiere a la interpretación de los tipos de posibilidad implicadas en el esquema general de la supervinencia, Chalmers distinguirá entre la supervinencia lógica y la supervinencia natural. De acuerdo con la primera, las propiedades B supervienen lógicamente sobre las propiedades A, si no hay dos situaciones lógicamente posibles indiscernibles en sus propiedades A que no sean indiscernibles respecto a sus propiedades B.⁷

La supervinencia natural entre dos tipos de propiedades es más débil que la lógica y se restringe a la correlación necesaria de propiedades en el mundo natural, de manera que las propiedades B supervienen naturalmente sobre las propiedades A si dos situaciones posibles que sean indiscernibles desde un punto de vista

⁷ Chalmers puntualiza que, en este caso, la supervinencia lógica no debe confundirse con la deducibilidad en un sistema de lógica formal, sino que se define en términos de mundos lógicamente posibles.

natural respecto a sus propiedades A, son también indiscernibles respecto a sus propiedades B. En principio, la supervinencia natural, según Chalmers, es demasiado débil para satisfacer los requisitos ontológicos de una posición materialista:

La supervinencia lógica y la natural tienen consecuencias bastante diferentes para la ontología; esto es, para la cuestión de lo que hay en el mundo. Si las propiedades B son supervinientes sobre las propiedades A, entonces hay un sentido en el que una vez los hechos A están dados, los hechos B se dan gratuitamente. Una vez Dios (hipotéticamente) aseguró que se dieran todos los hechos físicos en nuestro mundo, los hechos biológicos se dieron gratuitamente. Los hechos B simplemente redescubren lo descrito por los hechos A. Pueden ser hechos *diferentes* (un hecho acerca de los elefantes no es un hecho microfísico), pero no son hechos *adicionales*. Con la mera supervinencia natural, la ontología no es tan sencilla. Las conexiones legaliformes contingentes conectan distintas características del mundo. En general, si en nuestro mundo las propiedades B son meramente supervinientes naturalmente sobre las propiedades A, entonces *podría* haber un mundo en el que nuestros hechos A se dieran sin los hechos B. Como vimos antes, una vez Dios fijó todos los hechos A, para fijar los hechos B tuvo que trabajar más. Los hechos B son algo más allá de los hechos A y su satisfacción implica que hay algo nuevo en el mundo. (Chalmers 1996: 41)

Sin embargo, aun considerando que los hechos físicos del mundo, una vez fijados, determinan el resto de hechos, la relación de supervinencia no puede, por sí misma, cumplir con los requisitos mínimos de una posición fisicista. Incluso si hablamos de supervinencia lógica, en el sentido descrito más arriba, no está claro que a partir de ésta pueda definirse una posición fisicista mínima. Parece que el requisito de que la supervinencia se interprete como una relación de necesidad lógica excluye posiciones incompatibles con el fisicismo como el emergentismo o el dualismo; sin embargo, esta exclusión no es suficiente. En primer lugar, una de las intuiciones básicas del fisicismo, como señalé más arriba, es que, en cierto sentido, no hay hechos adicionales respecto a los hechos físicos y el requisito de la supervinencia lógica apunta en esta dirección. El problema es que la relación de supervinencia no

excluye que un conjunto de hechos del mundo que no sea físico determine, igualmente, el resto de hechos del mundo y, por ende, la totalidad de los hechos físicos. Un pluralista metafísico podría aceptar tranquilamente tal afirmación. Sin embargo, ésta es una posibilidad que un fisicista difícilmente estaría dispuesto a aceptar, ya que en tal caso los hechos físicos no tendrían prioridad ontológica respecto al resto de hechos del mundo.⁸ En segundo lugar, dado que el tipo de necesidad lógica a la que se apela es independiente de la deducibilidad en un sistema formal, la simple afirmación de que se mantiene una relación de supervinencia lógica entre todas las propiedades del mundo no físicas y las propiedades físicas no hace inteligible por sí misma la idea de que todo en el mundo dependa de lo físico y, sin una clarificación adicional, convierte la relación de supervinencia en una correlación bruta en el mejor de los casos. Estas insuficiencias han llevado a varios autores⁹ a rechazar la relación de supervinencia como la noción básica en torno a la cual articular el fisicismo, de modo que, aunque esta pueda constituir un criterio necesario de toda posición fisicista, no es una condición suficiente del mismo.

En cierto sentido, el hecho de que la relación de supervinencia sea compatible con diversas posiciones metafísicas facilita la respuesta a un tipo de objeción planteado a ciertas concepciones de fisicismo. Una de las razones que convierte en implausible la idea de que todo es físico es que la mera identidad es demasiado fuerte. Si queremos entender cómo es posible que una mesa sea algo físico, no podemos definir el término “mesa” según un patrón físico demasiado determinado, ya que una mesa puede estar hecha de múltiples materiales y puede ejemplificar una gran diversidad de propiedades. En el caso de la mente, es implausible considerar que un estado mental sea idéntico a un estado físico determinado o que solo una estructura biológica específica pueda ejemplificar dicho estado. Por otra parte, entendemos que un estado mental tiene determinados poderes causales en virtud de ciertas propiedades. Por ejemplo, saber que alguien tiene una determinada creencia nos permite llevar a cabo ciertas predicciones, sin

⁸ Schafer (2003) defiende que el fisicismo requiere, necesariamente, que haya un nivel fundamental o último de la realidad, y que éste sea el nivel de las propiedades físicas.

⁹ Véase Melnyk (2003) y Poland (1994).

necesidad de tener conocimiento alguno de física. Entender que S tiene una creencia equivale a entender que S tiene un estado mental que se relaciona de determinada manera con otros estados mentales, que es generado por determinados hechos y que da lugar a ciertos comportamientos. Si S cree que llueve, normalmente, podemos deducir que su creencia ha sido causada por el hecho de que llueve, que S saldrá a la calle con un paraguas, dado que probablemente no querrá mojarse, etc. Siguiendo este análisis, podríamos asumir que lo que define a un estado mental es el rol causal que desempeña. Una teoría de este tipo es lo que se conoce como funcionalismo. Ahora bien, un mismo rol causal puede ser desempeñado por diferentes realizadores, utilizando un término habitual en la bibliografía. Si una posición funcionalista fuera verdadera y se lograra identificar a los realizadores de cada propiedad funcional; y si además resultara que los realizadores de estas propiedades fueran físicos, en el sentido supuesto más arriba, no sólo se satisfaría la relación de supervinencia, sino que además habría una explicación de por qué esta se da, y se podría dotar de un contenido claro a la idea de que todo es físico.

Chalmers considera que el análisis funcionalista, en lo concerniente a lo mental, es exitoso, con la excepción de la experiencia consciente, como veremos. La compleción del proyecto funcionalista depende del éxito, en primer lugar, de un análisis conceptual que clarifique en términos de rol causal el ítem sujeto a reducción; en segundo lugar, cómo se determina cuáles son los realizadores de la propiedad funcional es una tarea llevada a cabo por la investigación empírica, de modo que no se puede conocer *a priori* la verdad del fisicismo. Por otra parte, el hecho de que una determinada propiedad física sea, o pueda ser, el realizador de una propiedad funcional depende, de acuerdo con el marco que se ha ido explicitando, de las leyes naturales del mundo. Por ejemplo, dado que la actividad neuronal depende de las propiedades químicas de determinadas sustancias, en un mundo en el que las leyes de la química fueran distintas a las del mundo real, sería posible que determinado tipo de actividad neuronal no pudiera cumplir con el rol

funcional que ejerce en nuestro mundo.¹⁰ En tal caso, la verdad del fisicismo sería contingente, ya que es necesaria la investigación empírica para determinar si una propiedad se realiza físicamente. Ahora bien, ¿puede conciliarse esta idea con la caracterización que Chalmers da del fisicismo en términos de supervinencia lógica?

Chalmers considera algunos ejemplos en los que las intuiciones fisicistas funcionan bien y sirven de contraste para mostrar la especificidad que plantea la experiencia consciente en la comprensión científica. Uno de los ejemplos es la identificación del calor con la agitación molecular media. Evidentemente, no sabemos a priori que el calor es agitación molecular; no obstante, parece que el descubrimiento de que el calor es agitación molecular agota, en cierto sentido, la posibilidad de plantear más interrogantes al respecto. Sabemos que el calor es un fenómeno generado por determinadas fuentes de energía, que produce la dilatación de los cuerpos, el aumento de la temperatura, etc. Si descubrimos que es la agitación molecular lo que es producido por la aplicación de una fuente apropiada de energía y que esta, a su vez, produce la dilatación de los cuerpos, el aumento de temperatura, etc., ¿no podemos decir que el calor es efectivamente la agitación molecular? Si todos los hechos referentes a la agitación molecular quedan fijados, ¿es posible la variación en los hechos que tienen que ver con el calor? Otro ejemplo elegido por Chalmers es el de la explicación de los procesos biológicos, como la reproducción o la vida. ¿Qué diferencia a un organismo vivo de la materia inerte? Un organismo vivo se adapta a su medio, se nutre, se reproduce y parece que un ser que lleve a cabo estas funciones es un ser vivo. La comprensión de cuáles son los procesos que realizan estas funciones se conoce en gran medida, como en el caso de los hechos químicos vinculados a la duplicación de la molécula de ADN, y los que tienen que ver con el desarrollo biológico de un ser vivo, y parece que, en principio, podrían determinarse los hechos físicos que subyacen a estos procesos. ¿Puede hablarse, pues, de unos hechos vitales adicionales respecto a los hechos físicos que se vinculan a estos? ¿Puede haber seres vivos sin los procesos mencionados o, al menos, sin procesos semejantes a éstos? ¿Sería posible que un

¹⁰ Esta forma de argumentar presupone la identidad de un determinado ítem con independencia de la variación de las leyes naturales de un determinado mundo. Sin embargo, ¿en qué sentido se puede hablar de *la misma actividad neuronal* si las leyes naturales han cambiado?

ser que al microscopio tiene la textura uniforme de una roca pudiera comportarse como un ser vivo? ¿Y en tal caso, lo consideraríamos un ser vivo?

Estos ejemplos apuntan a la idea de que una vez fijados los hechos físicos, ya no hay lugar para hechos adicionales. Ahora bien, ¿cómo puede determinarse a priori si un determinado tipo de hecho fija el resto de los hechos? ¿La fuente de estas intuiciones tiene alguna relación con nuestro conocimiento básico de los hechos físicos? Para poder determinar a priori si un determinado conjunto de hechos determina otro conjunto de hechos, debemos tener primero cierto conocimiento del contenido de términos como “calor” o “ser vivo”. Este contenido, determinado mediante el análisis funcionalista, sería el elemento a priori que requiere la concepción fisicista de Chalmers. Por otro lado, para llegar a saber cómo se realiza el rol funcional en cuestión, es necesaria la investigación empírica. Como veremos más adelante, Chalmers se servirá de un marco semántico mediante el cual tratará de excluir la necesidad metafísica como un tipo de necesidad distinta de la necesidad lógica. Con todo, no trataré dicha discusión en este capítulo. Más bien, me centraré en un movimiento dialéctico de la argumentación de Chalmers que, si bien es sutil, define la concepción metafísica que Chalmers tiene de la realidad y, por ende, por qué la experiencia consciente es una dimensión no reducible al resto de la realidad.

1.2. Fisicismo y microfisicismo

Volviendo a la cita de más arriba, recordemos que la idea básica del fisicismo para Chalmers es que, aunque puede haber hechos distintos de los hechos físicos, estos no pueden ser hechos adicionales. Y el ejemplo del que se sirve es el siguiente: “Pueden ser hechos *diferentes* (un hecho acerca de los elefantes no es un hecho microfísico), pero no son hechos *adicionales*” (Chalmers 1996: 38). Es fácil entender que, si hablamos de elefantes, no estamos hablando de entidades que formen parte de la física; y tampoco estamos diciendo algo respecto a las partículas que la física reconoce. Por otra parte, Chalmers se sirve de un ejemplo gráfico de algo que claramente es físico. Ahora bien, ¿es lo mismo un hecho físico que un hecho

microfísico? Ya he mencionado antes la problemática de determinar qué es un hecho físico; no obstante, resultaría difícil excluir el movimiento de un cuerpo visible a simple vista como un hecho físico a pesar de que, de entrada, no se trata de un hecho microfísico. Este movimiento dialéctico no es casual, sino que es bastante frecuente en la argumentación de Chalmers, y en ocasiones hace referencia a propiedades de un nivel inferior, o incluso, a propiedades fundamentales.¹¹

Son varios los filósofos que han señalado la diferencia entre una tesis fisicista y una tesis microfisicista (Kim 1998; Melnyk 2003; Hüttemann 2004). Lo que caracteriza el microfisicismo es una concepción jerárquica de la realidad, según la cual hay diferentes niveles en la realidad, de manera que las propiedades de un determinado nivel se realizan en propiedades de un orden menor. Lo que determina a qué orden pertenece cada nivel consiste en que, en la descripción de un nivel inmediatamente superior a otro, solo pueden aparecer partes constituyentes de este último y las relaciones entre estas. Por ejemplo, dado un conjunto de propiedades P perteneciente al nivel s, si las propiedades del tipo P son realizables, esto es, no se trata de un conjunto de propiedades básicas o fundamentales, el conjunto F de propiedades que las realiza deberá pertenecer a un nivel menor que s.

Un fisicista no microfisicista puede considerar que los hechos físicos determinan el resto de hechos del mundo, pero no tiene por qué aceptar que los hechos microfísicos determinan los hechos macrofísicos, sobre todo teniendo en cuenta que no existe una teoría física que dé cuenta unitariamente de lo macroscópico y lo microscópico. Ahora bien, esta diferencia es muy importante, ya que la fuente de la intuición que hace plausible cada una de estas posiciones es distinta. En el caso de un fisicismo microfisicista, su plausibilidad deriva, en gran parte, de la idea de que la física, concretamente la física de partículas, da una descripción de la realidad que tiene en cuenta sus partes más pequeñas. En este sentido, la plausibilidad del fisicismo es parasitaria de la idea de que si describimos

¹¹ En lo que sigue, el habla de propiedades físicas se restringe implícitamente a la clase de las propiedades fundamentales, a menos que se indique otra cosa. A veces hablaré de propiedades “microfísicas” o de “bajo nivel” (Chalmers 1996: 33).

Podemos imaginar que una vez se le den a un superser hipotético—digamos, el demonio de Laplace—todos los hechos microfísicos, podría “leer al completo” todos los hechos biológicos con dicha base (Chalmers 1996: 35).

las partes que forman un todo y sus relaciones, estamos ofreciendo una descripción metafísica adecuada del mundo o, al menos, nos aproximamos mejor al ideal de dicha descripción; mientras que el resto de descripciones del mundo, entre ellas las descripciones que tienen que ver con lo macroscópico, a lo sumo constituyen paráfrasis útiles en la práctica. Y es precisamente la física, concretamente la física de partículas, la que ofrece la descripción adecuada de lo más pequeño.

En el caso del calor o de la vida, al observar que, en última instancia, todo lo que hay son moléculas relacionándose entre sí de modos complejos, parece que ya no tiene sentido preguntar por algo más. La pertinencia de la pregunta por los hechos adicionales a los hechos del calor o de la vida sería semejante a la mostrada en el siguiente ejemplo: imaginemos que desde lejos vemos el trazo de una línea negra en una pared blanca y que al acercarnos lo suficiente percibimos una fila de hormigas moviéndose incesantemente. ¿Tiene sentido preguntar por un hecho adicional, en el sentido especificado más arriba, respecto al movimiento de las hormigas y a su disposición? La conclusión a la que llegaríamos es que no hay nada más que hormigas moviéndose; si éstas desaparecen, desaparece la línea que forman, si las hormigas cambian su ruta, cambiará igualmente la forma de la línea, y a cada disposición de las hormigas, corresponderá una única línea.¹²

Un fisicismo no microfisicista no puede servirse de esta imagen (o a lo sumo no puede ser la intuición que respalde su argumentación básica), aunque se encuentra libre de defender la concepción jerárquica de la realidad descrita más arriba.¹³ Chalmers no lleva a cabo ninguna distinción entre fisicismo y microfisicismo. No parece que esta sea una cuestión relevante en su argumentación

¹² En este último caso, se apela a un observador y por tanto para que lo dicho sea verdad hay ciertas condiciones que tienen que mantenerse estables (que el observador no padezca de una afección ocular, que haya luz, etc.). Como veremos más adelante, la necesidad de apelar a dichas circunstancias contextuales entraña una serie de dificultades que hacen difícil sostener este tipo de imagen.

¹³ Es importante señalar que no es lo mismo distinguir entre el contenido de dos tesis, en este caso el fisicismo y el microfisicismo, que negar que cada una de ellas sólo sea plausible a la luz de la otra, respectivamente. Si esto último no se hace, quizá no sea demasiado interesante apuntar a la diferencia entre ambas tesis. En el caso de Andrew Melnyk, a pesar de defender un fisicismo no microfisicista, al tratar la cuestión del reduccionismo en su posición, se sirve de la misma intuición microfisicista: “¿Realmente pensamos que el hecho de que un muro hecho de piedra en seco deba su existencia por completo a ciertas piedras dispuestas adecuadamente, en algún sentido, pone en duda la realidad del muro, sea en absoluto o en comparación con la de sus piedras constituyentes?” (Melnyk 2003: 75).

y, probablemente, el elemento importante en el problema de la conciencia no lo juega el fisicismo en sí, sino el modelo metafísico jerárquico de la realidad en el que lo físico encaja de un modo especial. Parece que las propiedades físicas son básicas en cierto sentido, o al menos son las propiedades que se encuentran en el nivel más bajo que hasta el momento conocemos. De hecho, ni siquiera importa tanto en qué consistan las propiedades físicas. Una confirmación de este diagnóstico es que en la discusión actual cada vez se da menos importancia al contenido concreto de lo físico, de manera que lo importante es no perder de vista que hay una asimetría entre el mundo de lo no mental y el mundo de lo mental, y una posición materialista tratará de mostrar que esta asimetría es espuria.¹⁴

1.3. Explicación reductiva e iluminación

En la sección anterior, la discusión en torno al marco general del fisicismo nos ha conducido a la delimitación de los rasgos generales de la concepción metafísica que subyace al mismo. En esta sección, trataré con más detalle el reflejo de esta concepción en su vertiente explicativa y en su relación con la concepción de Chalmers respecto a la idea de explicación reductiva. De un modo general, Chalmers define la explicación reductiva a la luz de las consideraciones metafísicas fundadas en la supervinencia:

Un fenómeno natural es explicable reductivamente en términos de algunas propiedades de nivel inferior si la propiedad de instanciar dicho fenómeno es superviniente lógica y globalmente sobre las propiedades bajas en cuestión. Un fenómeno es explicable reductivamente *simpliciter* si la propiedad de ejemplificar

¹⁴ Podemos encontrar dos ejemplos notables que reflejan esta manera de entender el materialismo: De este modo, una manera de entender “físico” simplemente sería “identificable como no mental—esto es, como representando propiedades que pueden identificarse independientemente de este aparato conceptual específicamente mental (Papineau 2002: 41). ¿Qué hay acerca de las mesas y las sillas que las hace ejemplos paradigmáticos de lo físico? Según Chomsky, es su volumen que ocupen espacios euclídeos, su mecánica de contacto, todas las cosas que ahora creemos que no son literalmente verdaderas de ellas (o al menos de sus constituyentes últimos). Diría, más bien, que es su no mentalidad (Levine 2001: 20).

dicho fenómeno es superviniente global y lógicamente sobre propiedades físicas.
(Chalmers 1996: 48)

La necesidad de que se cumpla esta condición para que sea posible una explicación reductiva parece evidente, al menos en cierto sentido de reducción, en la medida en que, si un fenómeno no superviniera lógicamente sobre las propiedades respecto a las que el fenómeno se va a reducir, siempre podría hacerse la pregunta de por qué el fenómeno de un nivel más bajo es acompañado por el fenómeno de orden superior. Sin embargo, el mismo Chalmers reconoce que esta noción estipulativa de explicación reductiva podría ser demasiado débil por dos razones. En primer lugar, los hechos de orden inferior podrían constituir un conjunto arbitrario de detalles sin una unidad explicativa. Ya avancé este problema cuando señalamos que, si una posición fisicista reposa exclusivamente en la noción de supervinencia, esta no se puede distinguir de posiciones tan paradigmáticamente no fisicistas como determinados tipos de dualismo. Si hay una intuición poderosa de que un tipo de fenómeno superviene sobre otro tipo de fenómeno, dicha relación debería tener algún fundamento inteligible; de otro modo, es arbitraria.¹⁵ En segundo lugar, teniendo en cuenta que distintas instancias de un mismo tipo de fenómeno pueden ser realizadas por distintos conjuntos de hechos, este tipo de explicaciones reductivas no puede explicar el fenómeno objeto de reducción como un tipo. Sin embargo, aun aceptando estas reservas, Chalmers considera, por una parte, que todo lo que necesita para sus argumentos es sostener que este tipo de explicación es una condición necesaria de una explicación reductiva. En tal caso, como veremos más adelante, un argumento que muestre que no es posible una explicación de tal tipo para un fenómeno, *a fortiori* será un argumento contra cualquier tipo de explicación reductiva de dicho fenómeno.

¹⁵ Chalmers no enfatiza, en este caso, la idea de que la evidencia empírica apunte a la realización física del resto de fenómenos no físicos, y que ese sea el fundamento de nuestras intuiciones respecto a la supervinencia. Como veremos en este mismo capítulo, hay una vía estándar para fundamentar el fisicismo que puede tolerar perfectamente la aparente arbitrariedad entre *explanans* y *explanandum*, pero Chalmers no se sirve de ella.

Por otra parte, Chalmers considera que hay un sentido útil de explicación reductiva que puede articularse en torno a la noción de supervinencia. En este sentido, Chalmers distingue entre explicaciones iluminadoras y explicaciones eliminadoras de misterios (*mystery-removing*):

La explicación reductiva no es lo único importante de la explicación. Su papel principal es eliminar cualquier sentido profundo de misterio que rodee un fenómeno de nivel alto. Y lo hace reduciendo la primitividad y la arbitrariedad del fenómeno en cuestión a la primitividad y arbitrariedad de procesos de niveles inferiores. En la medida en que los procesos de un nivel bajo pueden ser, en sí mismos, bastante primitivos y arbitrarios, una explicación reductiva quizá no nos dé un entendimiento profundo de un fenómeno, pero al menos elimina cualquier sentido de que algo adicional está sucediendo. (Chalmers 1996: 49)

Uno de los problemas de esta distinción es que no está muy claro en qué medida trazarla no es más que un acto adicional de estipulación que no resuelve el problema de la arbitrariedad; en última instancia, si los hechos a los que vamos a reducir un fenómeno determinado nos parecen arbitrarios, ¿en qué sentido hemos logrado una explicación reductiva? Pensemos en el ejemplo de la línea de hormigas: ¿nos parece arbitrario que las hormigas que trepan por la pared sean la línea que divisamos a lo lejos? Una explicación de por qué hay una línea negra en la pared parece que se agota en el hecho de que unas hormigas están trepando por la pared, y no tiene nada de arbitrario.

Sin embargo, no es cierto que una buena explicación de un hecho no nos pueda parecer arbitraria en cierto sentido. Imaginemos, por ejemplo, una investigación policial que trata de dar con un asesino y que, dadas las circunstancias, está muy claro que el asesino solo puede ser una persona de entre un grupo de sospechosos. Podemos imaginar que el asesino es A, y que ha sido especialmente escrupuloso y eficaz borrando cualquier rastro que pueda indicar tanto que él haya sido el asesino como su *modus operandi*. Pero todos los sospechosos tienen una coartada cierta y verificable, con la excepción de A. ¿No sería ésta una buena explicación de por qué A es el asesino, a pesar de que no se

pueda reconstruir una cadena de hechos que permita identificar a A como el ocupante del rol de asesino?¹⁶ El arte de la prestidigitación juega, precisamente, con el contraste que se genera entre la certeza de que hay una explicación rutinaria y satisfactoria de la ilusión que presencia el espectador, y la ausencia de la evidencia que permite reconstruir dicha explicación.

En cierta manera, uno de los argumentos más extendidos que se usa para defender el fisicismo tiene una estructura parecida. La idea básica es que, aunque no podamos saber exactamente de qué manera todo es físico, o de qué manera todo es físicamente realizable, podemos estar convencidos de que así es, dado que el nivel de la física goza de una autonomía causal que determina el resto de niveles, y que no deja otra opción sino el fisicismo. Este argumento se basa, o pretende basarse, en la evidencia empírica que proporciona la física, y se le ha llamado el argumento causal a favor del fisicismo.¹⁷ El argumento acepta, en primer lugar, que los estados mentales tienen efectos físicos, de manera que, por ejemplo, los estados mentales en los que alguien se encuentra cuando percibe un león causan los movimientos corporales de una conducta de huida. La segunda premisa, que se vincula con la evidencia que proporciona la física, afirma que todo efecto físico tiene una causa física antecedente; esta premisa tiene también una versión explicativa, según la cual todo efecto físico tiene una explicación física, o puede explicarse en términos exclusivamente físicos. Dadas estas dos premisas, todavía no se sigue una conclusión fisicista, ya que cabe la posibilidad de que lo mental y lo físico constituyan líneas causales independientes que convergen en un mismo efecto, pero que gozan de una identidad independiente. Sin embargo, resulta contraintuitivo aceptar como un fenómeno habitual y extendido que haya dos causas simultáneas de un mismo efecto¹⁸, y para ello se introduce la tercera premisa, según la cual los efectos físicos de los estados mentales conscientes no suelen estar determinados, simultáneamente, por dos causas distintas. Ahora bien, si se acepta que tanto los

¹⁶ Block y Stalnaker (1999) han señalado que las identidades o, al menos, determinado tipo de identidades no se explican. Ésta es una idea que discutiré más adelante.

¹⁷ Para una presentación del argumento ver Kim (1998).

¹⁸ Esto es lo que en la bibliografía se conoce como sobredeterminación masiva; un ejemplo es un fusilamiento, donde distintos impactos de bala producen la muerte, y aunque cada uno, por separado, no se hubiera producido, igualmente se habría causado la muerte.

estados mentales como los estados físicos de los que se sigue el efecto son eficaces causalmente, entonces el estado físico debe ser idéntico con el estado mental, y por lo tanto el fisicismo es verdadero. El argumento está construido en un nivel de abstracción tal que no se compromete con la identificación de la propiedad o conjunto de propiedades físicos que son idénticos con el estado mental, y ni siquiera necesita que esta identificación sea posible, dadas condiciones epistémicas ideales. Este fuerte realismo puede caracterizar, pues, a una posición fisicista. En cualquier caso, si esta fuera la mejor conclusión a la que se pudiera llegar, todavía persistiría un problema filosófico profundo en la medida en que se seguiría manteniendo una asimetría entre varios tipos de explicación; y no está claro en qué medida una barrera epistémica que impidiera determinar la relación precisa entre dos conjuntos de hechos no es un misterio.

Según Chalmers, la autonomía y la simplicidad de la física acercan una explicación de este tipo a una explicación iluminadora; por ejemplo, el mundo podría estar compuesto de leyes emergentes que gobernarán con su propia autonomía los niveles macroscópicos, o bien el comportamiento microfísico del mundo podría ser caótico y completamente carente de leyes, y en tal caso no habría manera de esperar una explicación reductiva iluminadora respecto a nada. Sin embargo, Chalmers no ofrece en ningún momento ninguna caracterización positiva de en qué consiste una explicación iluminadora, a pesar de que considera la autonomía y la simplicidad como rasgos de la misma. Su posición no va a necesitar la elaboración de dicha noción; ahora bien, las posibilidades que este planteamiento tiene en cuenta presuponen la concepción estratificada de la realidad que he descrito más arriba y quizá no deje espacio para otra explicación que no sea la que el mismo Chalmers ha definido como eliminadora de misterio. Si aceptamos la concepción estratificada de la realidad y la idea de que los niveles más bajos determinan el resto de niveles, ¿qué podría constituir una explicación iluminadora? La idea de que una explicación ilumina parece que encierra una referencia a un sujeto para el cual esta explicación es iluminadora. Una explicación de este tipo, por tanto, debería tener en cuenta las peculiaridades de un sujeto, para el cual un determinado fenómeno es misterioso u oscuro. Estas peculiaridades incluirían, no sólo el tipo de limitaciones

cognitivas, como las relacionadas con nuestros órganos sensoriales o nuestra memoria, sino también ciertos intereses que de un modo robusto se encuentran vinculados al tipo de investigación que llevamos a cabo. Ahora bien, ¿qué características generales debería tener la concepción que hemos descrito para poder tener en cuenta de las peculiaridades de los distintos sujetos que puedan aspirar a una explicación iluminadora? Volveré sobre este punto en la sección 1.5, pero antes esbozaré el mapa de posibles posiciones que, según Chalmers, pueden adoptarse ante el problema de la experiencia consciente.

1.4. La geografía de la discusión

La posición de Chalmers puede resumirse en dos tesis generales: (1) el fisicismo es falso, (2) la experiencia consciente superviene lógicamente o naturalmente sobre los hechos físicos del mundo. Juntas dan lugar a la posición que Chalmers (1996) denomina dualismo naturalista. El contenido de la primera tesis puede entenderse remitiéndonos a la caracterización de fisicismo presentada en este mismo capítulo: tal y como vimos, Chalmers define el fisicismo apelando a la noción de supervinencia, de manera que el fisicismo es verdadero si todos los hechos del mundo supervienen lógicamente sobre los hechos físicos. Sin embargo, no todos los hechos del mundo supervienen lógicamente sobre los hechos físicos; concretamente, los hechos vinculados a la experiencia consciente son hechos adicionales a los hechos físicos. La tesis (2) es una disyunción y puesto que la verdad de la tesis (1) implica que la experiencia consciente no superviene lógicamente sobre los hechos físicos del mundo, por silogismo disyuntivo, la experiencia consciente superviene naturalmente sobre los hechos físicos del mundo. Por otra parte, tal y como el calificativo de naturalista indica, Chalmers intenta eliminar cualquier connotación de misterio asociada a lo que tradicionalmente se ha entendido como dualismo, de manera que su posición no sólo sea compatible con la ciencia, sino que permita generar un marco en el que la experiencia consciente pueda ser explicada científicamente. Este segundo aspecto será tratado en el Capítulo 6. A continuación, analizaré brevemente la

caracterización que Chalmers ofrece de las posibles posiciones que pueden adoptarse frente a su argumento antirreduccionista y presentaré el mapa del debate. Para la exposición del esquema del debate tendré en cuenta, principalmente, Chalmers (2002).

Veámos que para que el materialismo fuera falso, era necesario que los hechos relativos a la experiencia consciente no supervinieran lógicamente sobre los hechos físicos. En principio, el establecimiento de una tesis de este tipo no depende del conocimiento empírico que tengamos de la conciencia. El tipo de necesidad lógica a la que se refiere Chalmers, como veíamos, no es la deducibilidad en un sistema formal, sino que hace referencia a la semántica de mundos posibles. Los argumentos que tratan de establecer la falsedad del materialismo mostrando la irreductibilidad de la experiencia consciente parten de una premisa epistémica que apunta a un hueco (*gap*) entre los hechos físicos y los hechos experienciales. Por ejemplo, parece que podemos concebir seres físicamente idénticos a nosotros que, no obstante, carecen de experiencia consciente (en la bibliografía al respecto son llamados zombis), o podemos concebir que los demás tienen una parte del espectro cromático invertido respecto a nosotros, de manera que no perciben los mismos colores, aunque no podamos llegar a saberlo. Pero los argumentos, como veremos en el Capítulo 3, no se limitan a constatar una limitación epistémica, sino que una consecuencia de que exista dicho hueco epistémico es que también hay ciertas consecuencias ontológicas respecto a la experiencia consciente, de manera que la conciencia no superviene sobre los hechos físicos y el materialismo es falso. El esquema general del argumento se puede sintetizar así:

- (1) Hay un hueco epistémico entre las verdades físicas y las verdades relativas a la experiencia consciente.
- (2) Si hay un hueco epistémico entre las verdades físicas y las verdades relativas a la experiencia consciente, entonces hay un hueco ontológico y el materialismo es falso.

(3) El materialismo es falso.

Dependiendo de cuál sea la posición frente al argumento, Chalmers distingue varias formas de comprender la experiencia consciente. Según Chalmers, hay tres tipos de materialismo. Los materialistas de tipo A niegan (1), de modo que niegan la existencia del hueco epistémico y afirman que, por ejemplo, no es concebible que seres físicamente idénticos a nosotros carezcan de experiencias conscientes. Una vez se hayan explicado las distintas funciones cognitivas y conductuales en relación a la mente, ya no hay nada más que explicar; Dretske (1995) y Lewis (1998) son representantes de este materialismo. Chalmers también incluye entre los materialistas de tipo A a los eliminacionistas, que niegan que haya experiencia consciente en el sentido relevante que genera el hueco epistémico descrito; este es el caso de Dennett (1991) y Rey (1995).¹⁹ En segundo lugar, los materialistas de tipo B reconocen la existencia de un hueco epistémico, pero niegan (2), de manera que no reconocen la existencia de un hueco ontológico. Aunque haya posibilidades extravagantes respecto a la experiencia consciente que sean concebibles, esto no quiere decir que sean metafísicamente posibles, del mismo modo que, en cierto sentido, podemos concebir que el agua no sea H₂O, ya que “agua” y “H₂O” son conceptos diferentes, pero no es posible que el agua no sea H₂O, y esto es algo que sabemos empíricamente. Los materialistas de tipo B pueden reconocer la existencia de problemas epistémicos interesantes en relación a la experiencia consciente sin que esto implique la falsedad del fisicismo. Chalmers incluye en este grupo, entre otros, a Block y Stalnaker (1999), Hill (1997), Levine (1983), Loar (1990/1997), Papineau (1993), Perry (2001) y Tye (1995). Finalmente, el materialismo de tipo C reconoce la existencia de un hueco epistémico, pero afirma que en algún momento se cerrará. Los materialistas de tipo C tienen en cuenta

¹⁹ En realidad, niegan el sentido fenoménico de la experiencia consciente que presentaré con más detalle en el Capítulo 2. Para Chalmers, la diferencia entre los eliminacionistas y los otros materialistas de tipo A es meramente terminológica, en la medida en que los eliminacionistas se refieren al sentido fenoménico de la experiencia como conciencia y niegan su existencia, mientras que los otros materialistas sólo reservan el término “conciencia” para las funciones cognitivas y conductuales relativas a la misma.

ejemplos en los que, en algún momento de la historia, no se disponía de los conceptos determinados para comprender un determinado fenómeno. Por ejemplo, siguiendo a Nagel (1974), los presocráticos no disponían de los conceptos necesarios para comprender que la materia es energía. Del mismo modo, como señala Stoljar (2007), quizá en este momento desconocemos un hecho relevante que, de ser conocido, ya no nos resultaría concebible que los zombis, por ejemplo, existieran. McGinn (1989) apunta a una limitación cognitiva por nuestra parte, si bien hay una solución, en principio, para el problema de la conciencia. La posición de Chalmers respecto al materialismo de tipo C es que no es una posición genuina, sino que, en caso de ser viable, resultará ser uno de los materialismos tipo A o B, o bien una de las posiciones antimaterialistas que se presentan a continuación.

En el otro grupo general de opciones, se acepta el esquema general del argumento antimaterialista y lo que distingue a cada posición son los términos en los que concibe dicho antimaterialismo. En primer lugar, el dualismo de tipo D se distingue por afirmar que en el nivel microfísico pueden intervenir causas no previstas en dicho nivel y la experiencia consciente es causalmente eficaz en el mundo físico; se trata de la posición que más comúnmente se conoce como interaccionismo. El dualismo sustancial de Descartes es el ejemplo más claro de la posición, aunque en este grupo también puede incluirse un dualismo de propiedades que reconozca la existencia de una sola sustancia. Ejemplos de interaccionismo son Foster (1991) y Popper y Eccles (1977). El dualismo de tipo E afirma que las propiedades físicas son ontológicamente distintas de las propiedades relativas a la experiencia consciente, pero niega la eficacia causal de estas últimas; esta posición es el epifenomenismo. También en el caso de este dualismo, puede darse en una versión en la que se distinguen dos sustancias o bien un dualismo de propiedades. En Chalmers (1996) se argumenta a favor de un dualismo de propiedades que él mismo denomina dualismo naturalista. Si bien desde el punto de vista metafísico es un dualismo, ya que la experiencia consciente es irreductible a lo físico, Chalmers considera que esto no debe implicar ningún tipo de misterio, en la medida en que la experiencia consciente se rige por leyes y está correlacionada con lo físico de un modo natural, de ahí que se califique como

naturalista. Puesto que en esta posición la experiencia consciente superviene nomológica o naturalmente sobre los hechos físicos, y no lógicamente, la experiencia consciente no tiene ningún efecto en el mundo físico o, al menos, un mundo físicamente idéntico al nuestro podría carecer de experiencia consciente²⁰. En este sentido, puede decirse que el dualismo de tipo E es la posición que Chalmers defiende en *The Conscious Mind* (1996). Finalmente, hay una posición que merece una mención aparte y que tiene un carácter más especulativo que las otras; se trata del monismo de tipo F. En esta posición, se especula con la idea de que la física se limita a estudiar las relaciones entre propiedades físicas o entre las propiedades físicas y nosotros, pero desconoce la base categórica de las propiedades disposicionales que estudia la física. La base categórica de las propiedades físicas, continúa la especulación, podrían ser las propiedades de la experiencia consciente, también llamadas fenoménicas o protofenoménicas, y, de hecho, afirma Chalmers, las propiedades de la experiencia consciente son las únicas propiedades intrínsecas que conocemos. Esta posición fue defendida por Russell (1927) por lo que también Chalmers la llama monismo russelliano. Si bien podría decirse que esta posición es una forma de fisicismo, ya que en última instancia habla de las bases categóricas de las propiedades físicas, el tipo de propiedades de las que habla esta teoría son muy diferentes a lo que actualmente se puede entender por una propiedad física, manteniendo cierta dualidad que la acerca a un dualismo.

Resumiendo, podemos distinguir, según este análisis, tres grandes tipos de posiciones. En primer lugar, el materialismo de tipo A: se trata de posiciones que se caracterizan por defender la supervinencia de la experiencia consciente sobre lo físico, en la medida en que ésta exista, de manera que tanto las posiciones reduccionistas como las eliminacionistas se incluyen en este grupo. En segundo lugar, en el materialismo de tipo B, podemos distinguir las posiciones que se comprometen con la verdad del fisicismo, si bien aceptan los argumentos antirreduccionistas; por lo que concluyen que los hechos físicos no se vinculan a

²⁰ En el Capítulo 4 trataré con más detalle la cuestión del epifenomenismo.

priori con las propiedades fenoménicas.²¹ El materialismo de tipo C no es, según Chalmers, una posición esencialmente diferente de un materialismo de tipo A o de alguna de las posiciones antimaterialistas, ya que en algún momento terminaría siendo un materialismo de tipo A, o bien un materialismo de tipo B. Finalmente, hay un tercer grupo de posiciones que niegan tanto el fisicismo como la supervinencia de la conciencia sobre lo físico y que incluye los dualismos de tipo D y E, y el monismo de tipo F.

Chalmers responde al materialismo de tipo B, que reconoce la existencia de necesidades metafísicas brutas, mediante el desarrollo de un modelo semántico que presentaré en el Capítulo 3. No obstante, el marco semántico presentado por Chalmers no podrá proporcionarnos por sí mismo el tipo de compromiso metafísico substantivo que es necesario para garantizar la irreductibilidad de la experiencia consciente. El materialismo de tipo A incluye tanto posiciones reduccionistas como eliminacionistas de un modo que borra, como veremos, una distinción fundamental entre ambas. En el caso de las posiciones reduccionistas, el funcionalismo es, para Chalmers, la vía más prometedora. Sin embargo, los argumentos antirreduccionistas de Chalmers, en caso de ser correctos, muestran que el análisis funcional no recoge el sentido fenoménico de experiencia consciente.²² El análisis funcionalista cometería un error de partida al no reconocer el carácter específico del aspecto fenoménico de nuestra experiencia, de modo que el tipo de vínculos causales que constituyen las propiedades funcionales no puede dar cuenta de la relación íntima y básica que mantenemos con nuestros propios estados mentales conscientes. Por otra parte, el eliminacionismo sólo se diferenciaría de este tipo de reduccionismo en la medida en que no reconoce la existencia de un objeto de explicación que se merezca el nombre de “experiencia”.²³ En general, la respuesta de Chalmers al grupo de posiciones que considera eliminacionistas será que, precisamente, por negar nuestras intuiciones más básicas respecto a la experiencia

²¹ No está claro que sostener la verdad del fisicismo y negar la reducibilidad de la experiencia consciente sea una posición coherente. Para ver algunas de las dificultades de este tipo de posición véase Kim (1998).

²² La presentación del sentido fenoménico de conciencia como distinto del sentido psicológico parece excluir, de entrada, el éxito de la reducción fenoménica, de modo que no es obvio en qué medida los argumentos de Chalmers no son sino redundantes respecto a las intuiciones de las que parte.

²³ Véase Chalmers (1996: 149).

consciente, constituyen una posición demasiado implausible como para ser tenida en cuenta.²⁴

En este sentido, un claro ejemplo de materialismo de tipo A que no escapará a este diagnóstico es la concepción de la experiencia consciente de Daniel Dennett articulada de un modo más extenso en *Consciousness Explained* (1991). Aunque el título de la obra apunta a un tipo de materialismo reduccionista, Dennett ha sido considerado como un representante paradigmático del eliminacionismo²⁵. Por otra parte, como ya he señalado, la inclusión de los materialistas reduccionistas y de las llamadas posiciones eliminacionistas en un mismo tipo de posición deja de lado diferencias relevantes, de manera que ignora un tipo de respuesta que no puede ser simplemente rechazada por las mismas razones que el materialismo reduccionista. En el Capítulo 5 presentaré una línea de argumentación contra la concepción de la experiencia de Chalmers que se servirá de la crítica que Daniel Dennett lleva a cabo de algunos supuestos metafísicos de la discusión sobre el problema de la conciencia. Esta línea argumentativa podría hacer pensar que la posición que se va a defender más adelante es un materialismo de tipo A y, concretamente, una posición eliminacionista. Sin embargo, como ya planteé en la introducción, uno de los objetivos principales de este trabajo es hacer explícitos los supuestos del debate sobre la experiencia consciente con el fin de mostrar una forma alternativa de acercarse a los problemas que se discuten en el seno de dicho debate, sin que se vayan a aceptar, por tanto, los términos del debate.

1.5. La concepción absoluta de la realidad, explicaciones desenmascaradoras y la noción de “dar sentido”

No está claro cómo el proyecto que anima el fisicismo debería posicionarse respecto a las distintas concepciones que hay del mundo y, por ende, en qué medida debería determinar el estatus de las peculiaridades de los seres cognitivos que adoptan una

²⁴ Según Chalmers: “El eliminacionismo respecto a la experiencia consciente es una posición poco razonable *solo* por nuestra familiaridad con ésta” (Chalmers 1996: 90), y “Para mí, el funcionalismo reductivo y el eliminacionismo parecen tan claramente falsos que encuentro difícil comprender cómo alguien podría aceptar una posición Tipo A” (Chalmers 1996: 150).

²⁵ Véase Seager (1999).

perspectiva respecto a la realidad. Una posibilidad es que, como parte de esas peculiaridades, haya creencias y representaciones del mundo que no puedan ser subsumidas en la descripción de la física, de manera que dichas creencias tendrán que ser explicadas como sesgos o prejuicios que no representan el mundo tal y como es. Un proyecto semejante es lo que Bernard Williams denominó la concepción absoluta de la realidad:

La concepción absoluta debería explicar, o al menos hacer posible que se explique, cómo pueden darse las representaciones más locales del mundo — y es esto lo que nos permitiría relacionarlas con cada una de ellas, así como con el mundo tal y como es independientemente de estas. Por ejemplo, debería permitirnos entender cómo ciertas cosas pueden parecernos verdes a nosotros, pero no a otros. Además, esta concepción del mundo debe posibilitar la explicación de cómo ella misma puede existir. (Williams 1978: 230)

Para hacer plausible esta concepción, podemos partir de la idea de que las peculiaridades del sujeto cognoscente en muchas ocasiones le impiden a este tener una visión de la realidad tal y como es. Nuestro sistema conceptual es muy rico, aunque parte de esta riqueza es el resultado, más bien arbitrario, de ser el tipo de seres que somos. Es evidente que, en gran parte, nuestros conceptos se encuentran ligados a circunstancias históricas y culturales que claramente podrían haber sido de otra manera; pero incluso conceptos que podríamos considerar universales, probablemente tienen un origen arbitrario. Nuestros órganos sensoriales y nuestro cerebro son el producto contingente de la selección natural y, con seguridad, el tipo de conceptos que podemos formar se encuentra constreñido por nuestra biología y, lo que es peor, probablemente dichos conceptos no se corresponden con la realidad tal y como es con independencia de nosotros.

Los colores son un ejemplo clásico en la historia de la filosofía de cuestionamiento de un tipo de concepto que es fundamental en nuestra percepción del mundo y que, no obstante, según la concepción dominante, no representa fielmente algo que existe en la realidad con independencia de nosotros. Ahora bien, para iniciar la búsqueda de esta concepción no podemos sino comenzar con los

conceptos de los que disponemos; ¿cómo podríamos, entonces, llegar a tener una concepción que represente el mundo con independencia de nosotros? Como Williams señala, si lo único que hacemos es tomar el conjunto de nuestras creencias como punto de partida, y tratamos de elegir aquellas creencias que son el efecto de cómo los objetos que reconocemos en el mundo nos afectan, lo único que haremos será repetir el conjunto de nuestras creencias. Por ejemplo, si decimos que nuestras creencias sobre los árboles son el resultado de la influencia que los árboles que hay en el mundo ejercen sobre nuestras creencias, y que por ello estas creencias representan el mundo tal y como es en sí, no hemos dicho nada sustantivo y tampoco hemos discriminado un subconjunto relevante de creencias. Por otra parte, si tratáramos de dar con una idea del mundo anterior a nuestras representaciones no determinaríamos, tampoco, ningún contenido concreto. No obstante, según Williams, existe una tercera posibilidad:

Esa posibilidad, como he explicado, depende fuertemente de la noción de explicación. La sustancia de la concepción absoluta (en contraposición a las ideas vacías e intangibles “del mundo” que se presentaron antes) radica en la idea de que podría explicar de un modo no vacío cómo es posible ella misma, así como las distintas visiones perspectivistas del mundo. Un rasgo importante de la ciencia moderna es su contribución a la explicación de cómo criaturas con nuestro origen y características pueden entender un mundo con las propiedades que esta misma ciencia atribuye al mundo. (Williams 1985: 139-140)

El criterio del que se sirve Williams para poder iniciar la búsqueda de la concepción absoluta de la realidad es la noción de explicación y, a la vez, vincula la concepción absoluta de la realidad a la ciencia moderna y a su poder explicativo. La historia de la ciencia moderna está llena de ejemplos en los que creencias extendidas o teorías ampliamente aceptadas son rechazadas a la luz de una nueva teoría con una gran fuerza explicativa. De entrada, el proyecto fisicista no es asimilable al proyecto de la concepción absoluta de la realidad, ya que una posición fisicista no tiene que cumplir con las mismas exigencias explicativas y no está claro en qué medida el fisicismo tendría que poder dar una explicación de las distintas concepciones locales

o antropocéntricas; de hecho, el argumento causal en favor del fisicismo no trata de dar una explicación de cómo los aspectos de la realidad aparentemente no reducibles a la física encajan en el marco fisicista. Este punto es importante, ya que habría que evaluar hasta qué punto un hueco explicativo puede ser un argumento contra el fisicismo y, como veremos, el argumento antifisicista de Chalmers recurre a los nexos entre concebibilidad y metafísica. El fisicismo comparte con la concepción absoluta de la realidad la intuición básica que la motiva: se trata de una descripción privilegiada de la realidad, y es privilegiada, precisamente, por su poder explicativo y, más concretamente, por el poder explicativo de la física. Por otra parte, si hubiese fenómenos de los que el fisicismo no pudiese dar cuenta, la posición filosófica resultante sería insatisfactoria al no poder explicar el origen de dicha limitación, y posiblemente la única vía coherente sería rechazar dichos fenómenos como ilusorios.

El fisicismo no es una teoría científica sino, en todo caso, una teoría metafísica que toma como evidencia el poder explicativo de la física. ¿Pero cómo puede la evidencia de la física llevarnos más allá de la física? Al fin y al cabo, si la evidencia de la física permite extraer conclusiones que la trascienden, ¿por qué los físicos no extraen esas conclusiones? ¿Por qué la ciencia física no se extiende para abarcar ese tipo de cuestiones? Paradójicamente, como he tratado de mostrar, el contenido concreto de la física no juega un papel demasiado importante en la discusión en torno al fisicismo. En última instancia, la pregunta por la relación entre lo físico y lo mental se reducía a una dicotomía bastante más abstracta entre lo mental y lo no mental; y una solución satisfactoria en este marco eliminaría las diferencias entre lo mental y lo no mental. Thomas Nagel, en *The View from Nowhere* (1983), tiene en cuenta una concepción semejante a la concepción absoluta de la realidad de Williams, y la denomina la concepción fisicista de la objetividad.²⁶ La dialéctica que conduce a este tipo de concepción bebe, en principio, de lo que la física nos enseña:

²⁶ Aunque esta concepción es compatible con el fisicismo, cuya verdad se seguiría si tal concepción llegara a formularse coherentemente, es una tesis que debe ser distinguida del fisicismo en la medida en que supone un modo general de entender la objetividad y es algo más que una conclusión extraída a partir de la evidencia empírica.

El desarrollo se da por fases, y cada una de éstas da una imagen más objetiva que la precedente. El primer paso es ver que nuestras percepciones son causadas por la acción de cosas sobre nosotros, a través de sus efectos en nuestros cuerpos, que son, en sí mismos, partes del mundo físico. El siguiente paso es darse cuenta de que, puesto que las mismas propiedades físicas que causan percepciones en nosotros a través de nuestros cuerpos también producen diferentes efectos en otras cosas físicas y pueden existir sin causar percepciones en absoluto, su naturaleza verdadera tiene que ser separable de su apariencia perceptiva y no tiene por qué parecersele. El tercer paso es intentar formar una concepción de esa naturaleza verdadera independiente de su apariencia bien sea para nosotros, bien para otro tipo de perceptores. Esto no sólo significa no pensar en el mundo físico desde nuestro propio punto de vista, sino no pensar en él, tampoco, desde un punto de vista perceptivo humano más general: no pensar en cómo se ve, se siente, huele, sabe o suena. De este modo, estas cualidades secundarias quedan fuera de nuestra imagen del mundo externo y se piensa estructuralmente en las cualidades primarias subyacentes como forma, tamaño, peso y movimiento. (Nagel 1983: 14)

Está claro que la idea de la existencia de una “naturaleza verdadera” (*true nature*) no forma parte de la física, al menos de la ciencia física que practican los científicos. Ahora bien, ¿es esta idea lo suficientemente neutral para poder figurar en el tipo de concepción que Nagel describe? El argumento para determinar la naturaleza verdadera de las propiedades físicas establece una serie de compromisos que difícilmente son neutrales. El argumento trata de concluir que, puesto que las propiedades físicas existen con independencia de que causen percepciones, su “naturaleza verdadera” es independiente de dichas percepciones. Pero este argumento no conduce por sí solo a la concepción absoluta de la realidad, sino que, en principio, solo identifica un rasgo específico de las propiedades físicas. Lo que hace especial a las propiedades físicas o no mentales es que el conjunto de redes causales en las que se encuentran incardinadas son independientes de lo mental, al menos en el sentido de que estas redes causales existen con independencia de que haya seres perceptores que, por efecto de determinadas propiedades físicas y dadas ciertas condiciones, llegan a tener determinadas percepciones. Por otra

parte, no está claro que la noción de causalidad pueda desempeñar el papel fundamental que se le atribuye, como una noción que va a permitir superar las distintas perspectivas locales. Si atendemos al concepto de causa que hasta ahora se ha empleado en la discusión, es evidente que no se trata de un concepto exclusivamente físico, en el sentido más técnico, ya que tiene en cuenta la causalidad entre lo físico y lo mental; ahora bien, el concepto preteórico de causa es dependiente de un contexto y relativo a un ser con unos intereses y una perspectiva. Por ejemplo, ¿cuál es la causa de una muerte por apuñalamiento, el apuñalamiento, la motivación del asesino, la pérdida de sangre de la víctima o la ausencia de afluencia de oxígeno y nutrientes que sufre el cerebro como resultado de la pérdida de sangre? La respuesta a estas preguntas es adecuada, en cada caso, en función de si las lleva a cabo un sociólogo, un criminólogo, un profesor universitario de medicina, un médico forense o una persona que lea la sección de sucesos de un diario. Y si el concepto de causalidad que se emplea como fundamento de la concepción absoluta de la realidad es preteórico, quizá habría que desarrollar un concepto más preciso o técnico, que no esté sometido a las limitaciones del concepto preteórico.²⁷

De entrada, que partamos de una perspectiva ya dada no supone un problema serio para la concepción absoluta de la realidad, ya que, como veíamos, no podemos sino partir de nuestras propias creencias y representaciones para iniciar su búsqueda. Pero si admitimos que, efectivamente, el criterio central que nos va a permitir discriminar un subconjunto de creencias como privilegiado es su poder explicativo y si el tipo de explicación al que nos referimos se sirve de una concepción preteórica de causalidad, no está muy claro en qué medida podríamos alcanzar la descripción privilegiada de la realidad que buscábamos, ya que dicho subconjunto de creencias se encontraría inextricablemente ligado a nuestra perspectiva regional. Del mismo modo, el carácter privilegiado del que gozaría una descripción fisicista adecuada de la realidad sería también dudoso. No hay que olvidar que, adicionalmente, no disponemos de una caracterización de “físico” que

²⁷ Para una crítica de la noción de causalidad en relación con el proyecto fisicista véase Corbí y Prades (2000). En el Capítulo 4 se tratará este tema en mayor detalle.

sea muy determinada, dado el propósito del fisicismo. En este sentido, el precio de que las nociones centrales que articulan el fisicismo no sean demasiado concretas es que tampoco está muy claro qué implica la verdad del fisicismo. Si todo se reduce a que el fisicismo excluye la posibilidad de mundos causalmente extravagantes, el fisicismo es compatible con muchas otras posiciones que son paradigmáticamente no fisicistas. En general, no hay ninguna posición filosófica viva y plausible que afirme que un objeto comienza a moverse por arte de magia, que en la cabeza de un alfiler hay cien almas encerradas, o que los objetos materiales desaparecen del mundo y vuelven a aparecer espontáneamente. Ahora bien, ¿cómo puede una concepción fisicista del mundo ser algo más que el rechazo general de estas posibilidades? En la misma línea, cuando Chalmers menciona la posibilidad de que el mundo sea un conjunto caótico de leyes emergentes, o bien que el nivel microfísico de la realidad se comporte de un modo arbitrario, ¿qué tipo de posibilidad está teniendo en cuenta? ¿Acaso puede seguir hablándose de nivel microfísico cuando todo lo que hay son conjuntos arbitrarios de hechos o regularidades inconexas? ¿En tal caso, tiene sentido seguir hablando de mundo, o del mismo mundo con diferentes leyes, de tal manera que las instancias de propiedades y objetos que se encuentran sometidos al desorden del mundo caótico conservan su identidad respecto a los ítems de nuestro mundo? En el Capítulo 4 analizaré con más detalle los problemas a los que la posición de Chalmers tiene que enfrentarse como consecuencia de sus planteamientos causales.

Hay otro aspecto que es importante analizar respecto a la idea de perspectiva. Aunque partir de la propia perspectiva es un paso necesario para alcanzar una descripción privilegiada de la realidad, hay un sentido en que apelar a una perspectiva particular es incompatible con la búsqueda de una concepción de cómo son las cosas. Este es, incluso, un lugar común asociado a la ciencia moderna. Antes de la revolución científica, era frecuente recurrir a ciertos prejuicios para extraer conclusiones científicas: por ejemplo, el número de esferas celestes que se asumía que componían el universo o la trayectoria de los planetas podía determinarse por supersticiones no cuestionadas o incluso por motivos estéticos, y no había un método definido que permitiera distinguir los prejuicios de los rasgos

objetivos de la realidad. Sin duda, el éxito de la ciencia moderna se explica, en parte, por la inhibición de todos los elementos subjetivos que podrían sesgar una investigación. Idealmente, no debería importar quién lleva a cabo una investigación científica si respeta ciertas reglas metodológicas y, como resultado de estas buenas prácticas epistémicas, podemos superar nuestros sesgos subjetivos y llegar a tener un conocimiento de la realidad independiente de nosotros. La historia de la ciencia ofrece numerosos ejemplos de cómo la correcta aplicación del método científico nos ha permitido rechazar concepciones erróneas de la realidad. La creencia de que el Sol se mueve alrededor de la Tierra, de que la Tierra es plana o la creencia en la existencia del flogisto o las hadas son algunos de esos ejemplos. Barry Stroud ha acuñado la expresión “explicación desenmascaradora” para referirse a este tipo de explicaciones, veamos en qué consiste:

Llamaré a una explicación que explica la apariencia de algo, o explica la creencia en ese algo, sin tener que suponer que dicha creencia sea verdadera, una explicación “desenmascaradora”. Esta revela que la base o la fuente de una creencia no está conectada del modo correcto con su verdad. De este modo, desenmascara o expone una creencia o apariencia por lo que es—una ilusión, una creencia falsa o una mera apariencia. (Stroud 1999: 75)

Stroud se sirve de un ejemplo sencillo para ilustrar este tipo de explicación. Imaginemos que alguien que lleva gafas de sol se encuentra frente a un muro blanco y cree que dicho muro es amarillo. La explicación desenmascaradora podría dar cuenta de la falsa creencia de que el muro es blanco apelando al efecto que produce el uso de las gafas de sol en la percepción del color y al olvido que el sujeto tiene de llevar las gafas. Si seguimos el ejemplo, podemos asumir que una explicación desenmascaradora tiene que cumplir, al menos, con tres requisitos. En primer lugar, tiene que haber un argumento o hay que apuntar a un hecho que muestre que la creencia que va a ser desenmascarada es falsa (o es una apariencia o una ilusión), en este caso, podríamos simplemente ver que el muro es blanco. En segundo lugar, hay que mostrar por qué la fuente de la creencia no está conectada de un modo correcto con la verdad de dicha creencia; en el ejemplo, el uso de las gafas de sol

junto con el olvido de las mismas explica por qué el sujeto tiene la creencia.²⁸ Finalmente, en la explicación no debe suponerse la verdad de la creencia que se pretende desenmascarar.

Por supuesto, una explicación desenmascaradora necesita de un *explanandum* y un *explanans*. En el caso del muro, no hay nada problemático en identificar la creencia de que el muro es amarillo como el *explanandum* y, en general, podemos encontrar claros ejemplos de explicaciones desenmascaradoras en la vida cotidiana. Sin embargo, según Stroud, es muy difícil para la concepción absoluta de la realidad dar contenido a su *explanandum*, dado que tiene que dar cuenta del sumamente complejo conjunto de nuestras creencias a partir de la austera concepción de la física. El fisicismo, unido a la concepción funcionalista esbozada más arriba, trata de reducir lo mental al rol causal desempeñado por un determinado realizador físico, de manera que trata de mostrar que lo mental es una instancia más de lo físico. Pero para poder llevar a cabo el proyecto desenmascarador del fisicismo es necesario poder comprender el contenido del *explanandum*, en este caso lo mental, de manera que pueda decirse posteriormente que lo mental es físico. ¿Pero cómo podría hacerse esto si el lenguaje físico no incluye terminología mental? ¿Cómo podría darse cuenta, en primera instancia, de la complejidad de lo mental partiendo de la relativamente pobre imagen de la física? Las explicaciones desenmascaradoras de la vida cotidiana se llevan a cabo en el seno de un sistema de conceptos que no se cuestiona. Sin embargo, el proyecto fisicista trata de situarse más allá de la estructura conceptual cotidiana e incluso aspira a revisarla. El problema es que es necesario reconocer la existencia de lo mental, al menos en cierto sentido, para poder tener un *explanandum*, pero, al mismo tiempo, ha de poder mostrarse que lo mental no es lo que parece o que, incluso, carece de realidad. En este sentido, un problema básico al que el marco esbozado tiene que dar solución es dotar de contenido al *explanandum*, que en este caso serán los hechos mentales relativos a la experiencia consciente.

²⁸ Este segundo elemento es lógicamente independiente del primero. En el ejemplo dado, el muro podría ser amarillo, pero si la persona que percibe el muro lleva gafas de sol y se ha olvidado de que las lleva, el origen de su creencia no está conectado de un modo adecuado con su verdad.

Siguiendo a Stroud, hay otro desafío al que se tiene que enfrentar la concepción absoluta de la realidad que parece que tiene que ver de un modo más directo con la idea de que una explicación sea iluminadora. En una explicación desenmascaradora de lo mental hay que dar cuenta de la relación entre los hechos físicos y los hechos mentales, aunque sea para explicar el carácter ilusorio de estos últimos. Respecto a la relación entre lo físico y lo mental, no bastaría con proponer una reducción o una interdefinición de lo mental y lo físico para tener una explicación desenmascaradora, sino que es necesario tener una comprensión previa de qué contaría como un vínculo explicativo adecuado entre ambos tipos de hechos; en caso contrario, no habría manera de evaluar el éxito de la explicación. Una forma de entender esta explicación podría ser recurrir a leyes psicofísicas y, de hecho, esta será, en cierto modo, la solución propuesta por Chalmers. El problema es que las explicaciones de este tipo acaban reconociendo la existencia de una mera correlación bruta. Chalmers es consciente del límite de una explicación de este tipo, y forma parte de su definición de explicación reductiva la idea de que un tipo de arbitrariedad sea, simplemente, asimilable a un tipo de arbitrariedad más básico, en este caso, el de la física. También Stroud reconoce que, en última instancia, no podemos sino recurrir a correlaciones brutas en nuestras explicaciones. Sin embargo, parece claro que recurrir a una correlación bruta no constituye siempre una explicación satisfactoria. Stroud menciona la relación entre el mundo y nuestras percepciones del mismo como ejemplo: según Stroud, normalmente, la explicación de que percibamos un objeto cúbico es que tenemos un objeto cúbico delante de nosotros, y no el hecho de que haya una correlación entre objetos cúbicos y la percepción de objetos cúbicos. Dicho de otro modo, parece natural pensar que hay un vínculo inteligible entre nuestras percepciones y lo que percibimos. Si se parte de la visión de la física para alcanzar la concepción absoluta de la realidad, ¿cómo podrían vincularse inteligiblemente hechos físicos y hechos mentales? ¿Hay lugar para una explicación iluminadora en este proyecto?

Al final de su obra, Williams matiza su posición respecto a la concepción absoluta de la realidad e incluso afirma que el proyecto de esta concepción podría ser imposible de cumplir. En Williams (2006), al reflexionar sobre la contraposición

entre una concepción humanística de la filosofía y una visión científicista, Williams apunta a una falacia *non sequitur* en relación al rol que se asigna a la ciencia, consistente en afirmar que, si la ciencia describe el mundo tal y como es, ya no hay espacio para una actividad filosófica independiente de la ciencia. Según Williams, dicho argumento solo funcionaría bajo el supuesto de que la finalidad de cualquier actividad filosófica fuera describir el mundo como es en sí mismo.²⁹ Y si esto fuera así, quizá no habría una frontera definida entre la ciencia y la filosofía, y esta última se limitaría a anticipar algunos de los desarrollos científicos. Sin embargo, hay un tipo de reflexión que, si bien no es independiente de producir un discurso verdadero, no se limita a la verdad. Es lo que Williams denomina “dar sentido” (*making sense*):

Cuando tratamos de dar sentido a un suceso particular, a menudo contamos una historia sobre una secuencia de hechos que condujo al mismo. Si le damos sentido (o lo explicamos, o llegamos a entenderlo), debemos asumir que los elementos de la historia son verdaderos, pero esto, por supuesto, no es suficiente: la secuencia de hechos tiene que tener sentido para nosotros, y dar sentido al resultado. Dicha historia es un tipo de narración. (Williams 2002: 233)

Está claro que hay muchos tipos de narración y la relevancia de elegir una u otra se encuentra indisolublemente ligada a un público, un “nosotros”, para el que debe tener sentido lo narrado. Este “nosotros” podría estar constituido por los seres racionales, la humanidad, una nación, la comunidad científica, etc., y será ese “nosotros” lo que determine si una narración es adecuada o no. Según Williams, un tipo de narración puede ser la presentación de un proceso natural como la instancia de una regularidad o una ley. Esta sería una narración habitual en el ámbito de las ciencias naturales. Sin embargo, es una narración posible entre otras. Otro tipo de narración habitual es la que agrupa una serie de sucesos en relación a las intenciones de un agente: un detective que trate de comprender la muerte de la víctima de un asesinato difícilmente se dará por satisfecho con una comprensión de cómo el veneno suministrado a la víctima produjo un paro cardíaco, e incluso si

²⁹ Véase Williams (2006: 3).

desconoce el mecanismo de actuación del veneno, pero descubre que el asesinato fue un crimen pasional, su investigación podría llegar a su fin. Claramente, la noción de “dar sentido” no puede estar ligada a una narración única o privilegiada, en la medida en que las cosas que tienen sentido para un grupo no lo tienen para otro. Por otra parte, que algo tenga sentido para un grupo determinado dependerá de los intereses y las necesidades de dicho grupo.

Una pregunta fundamental que hay que plantear en relación al marco del fisicismo es en qué medida el tipo de narración que constituye el fisicismo puede dar cuenta de todos los aspectos que son relevantes para el ser humano. Esta pregunta es especialmente importante en relación a la experiencia consciente, ya que, como veremos, entendemos que la conciencia es lo que hace la vida digna de ser vivida, de manera que no es una dimensión respecto a la que no tengamos intereses fuertes o que nos importe poco. Sin embargo, Chalmers considera que su mismo modelo de explicación en relación a la conciencia no es iluminador; lejos de discutir el marco esbozado, la conclusión de Chalmers no será sino un intento de llevar a sus últimas consecuencias el proyecto de una concepción absoluta de la realidad, reconociendo la experiencia consciente como una parcela intratable por dicha concepción. No es extraño que Chalmers no pueda dar con una explicación iluminadora de la experiencia consciente. La idea de que algo sea iluminador parece próxima a que algo tenga sentido; sin embargo, el fisicismo presenta cierta descripción como privilegiada y al margen de cualquier interés humano, de manera que, difícilmente, podrá constituir una explicación iluminadora. En el Capítulo 6, mostraré que la concepción de la experiencia consciente de Chalmers, con independencia de que sea verdadera o falsa, carece de potencial para explicar por qué la experiencia consciente es tan importante para nosotros.

Recapitulando, presenté inicialmente el fisicismo como el marco general en el que se plantea la discusión dominante en filosofía de la mente. Chalmers toma la relación de supervinencia lógica como el elemento que articula su noción de fisicismo. No obstante, la relación de supervinencia es compatible con posiciones paradigmáticamente no fisicistas. Lo que explica la supervinencia lógica, según Chalmers, es que, en última instancia, la mayor parte de los estados mentales es

reducible a estados funcionales realizados por propiedades físicas. El modelo funcionalista de la mente, si fuera exitoso, permitiría, partiendo de definiciones funcionales de dichos estados y mediante la investigación empírica, dar con los realizadores físicos que satisfacen dicha definición. En última instancia, parece que el poder explicativo de esta imagen no proviene de la física, sino de una concepción metafísica que distingue una jerarquía de niveles en la realidad. Según Chalmers, solo con que consiguiéramos mostrar que todo lo mental superviene lógicamente sobre lo físico, dispondríamos de una explicación reductiva, eliminadora de misterio. Pero dicha explicación, según Chalmers, no nos daría una comprensión profunda de lo mental, no sería iluminadora.

Al indagar sobre qué podría constituir una explicación iluminadora, introduje la idea de perspectiva y planteé la cuestión de cómo un proyecto como el del fisicismo debería relacionarse con las distintas perspectivas respecto a fenómenos que, de entrada, no parecen físicos. La formulación más completa de un proyecto que trate de explicar las distintas perspectivas respecto a la realidad es la concepción absoluta de la realidad de Bernard Williams, la cual aspira a ser una concepción privilegiada de la realidad tal y como es en sí, explicando, a la vez, la existencia de las distintas perspectivas locales a partir del austero lenguaje de la física. El fisicismo comparte con la concepción absoluta de la realidad la aspiración de convertirse en una descripción privilegiada. Pero el éxito del proyecto tiene que poder responder a dos preguntas. En primer lugar, es necesario poder identificar lo mental a partir del punto de vista físico. En segundo lugar, desde el punto de vista fisicista, hay que poder dar cuenta de la relación entre lo mental y lo físico de manera que pueda entenderse cómo nuestros estados mentales responden al mundo de un modo inteligible. El tipo de actividad que en este caso llevaría a cabo la reflexión filosófica no se reduciría a decir algo verdadero, sino que se explicaría mejor apelando a la noción de Williams de “dar sentido”. Sin embargo, este es un movimiento que el proyecto fisicista tiene vedado, ya que la noción de “dar sentido” hace referencia a un público con unos intereses, y el fisicismo se presenta como una descripción de la realidad ajena a cualquier interés.

2. Los dos conceptos de lo mental: una distinción problemática

2.1. Una primera aproximación

La primera dificultad a la que se enfrenta un estudio de la experiencia consciente es la ambigüedad. Cuando hablamos de la conciencia podemos estar refiriéndonos a muchos fenómenos, y son numerosos los problemas planteados, incluso teniendo en cuenta una única acepción del término. David Chalmers señala cómo en muchos trabajos lo que se presenta como una respuesta al problema de la conciencia no es sino una respuesta a otros problemas que, sin carecer de interés, dejan de lado la cuestión que él mismo denomina el problema difícil (*hard problem*) de la conciencia. En ocasiones, la formulación del problema puede parecer superficialmente la misma, aunque en el fondo trate cosas tan dispares como una disputa dentro de una práctica cotidiana, una cuestión empírica, sujeta a un estudio científico convencional, o un enigma metafísico.

La pregunta básica que plantea *The Conscious Mind* (1996) es: “¿Cómo un sistema físico como el cerebro puede, a su vez, ser un *experimentador*? ¿Por qué debería haber algo *que se siente como ser*³⁰ tal sistema?” (Chalmers 1996: xi). Estas dos preguntas apuntan a la misma cuestión, de manera que la segunda califica a la primera dándole su cariz propiamente filosófico. El problema que Chalmers considera central remite al célebre artículo de Thomas Nagel *What is it like to be a bat?* (1974), en el cual se plantea la dificultad básica de poder acceder al conocimiento de la experiencia subjetiva de un ser vivo cuyo sistema perceptivo difiere notablemente del nuestro, como es el caso de un murciélago. Y este es, precisamente, este aspecto de la experiencia consciente el objeto de estudio de *The Conscious Mind* (1996):

La cuestión principal está mejor caracterizada como “la cualidad subjetiva de la experiencia”. Cuando percibimos, pensamos y actuamos, hay un zumbido de fondo

³⁰ El texto original dice: “Why should there be something it is like to be such a system?” Aunque en el texto original no aparezca el verbo “*to feel*”, equivalente al verbo “sentir” en castellano, la traducción se aproxima el significado original.

de causalidad y un procesamiento de la información, pero este procesamiento, normalmente, no se da en la oscuridad. También hay un aspecto interno; hay algo que se siente como ser un agente cognitivo. Este aspecto interno es la experiencia consciente. La experiencia consciente incluye desde las sensaciones vívidas de color hasta las experiencias de los aromas más tenues del ambiente; desde dolores agudos hasta la experiencia elusiva de tener un pensamiento en la punta de la lengua; desde sonidos y olores mundanos hasta la envolvente grandeza de la experiencia musical; desde la trivialidad de un picor molesto al peso de la profunda angustia existencial; desde la especificidad del sabor de la menta a la generalidad de la experiencia que uno tiene del yo. Todas estas cosas tienen una cualidad experimentada diferente. Todas son partes destacadas de la vida interna de la mente. (Chalmers 1996: 4)

Tal y como se caracteriza el problema, queda establecida una línea divisoria general entre dos aspectos de lo mental. Podemos hablar de un conjunto de procesos causales y de un procesamiento de información que se contraponen a un aspecto “interno”, este último coincidente con la experiencia consciente, con la “vida interna” de la mente. La utilización de estos calificativos es claramente metafórica, en la medida en que la distinción, en un sentido estricto, no remite a una distinción espacial, por ejemplo, entre procesos que ocurren en el cuerpo y procesos que ocurren fuera del mismo. En este sentido, el procesamiento de la información por parte del cerebro es interno. La distinción no coincide tampoco con la que hay entre estados que remiten a cosas que ocurren en nuestro cuerpo, como un dolor, y estados que remiten a lo que ocurre fuera de nuestro cuerpo, como la experiencia de color que tenemos al mirar una manzana, ya que ambos estados pertenecen al aspecto subjetivo de la experiencia. La referencia a procesos que ocurren en la oscuridad tampoco puede entenderse literalmente. Hablar de “oscuridad” en contraposición, quizá, a la luz aportada por la experiencia consciente señala el aspecto adicional de los procesos que no ocurren inconscientemente; la metáfora puede reflejar cómo la experiencia visual constituye el paradigma de la experiencia consciente, aunque claramente no se limita a esta.

En cualquier caso, aunque esta caracterización no fija con precisión en qué consiste la experiencia consciente, el mismo Chalmers sostiene que a este respecto todo lo que se puede hacer es ofrecer ejemplos.³¹ Parece que la idea de “lo interno” de la experiencia consciente debe ser entendida de un modo especial para llevar a cabo una distinción filosóficamente relevante. ¿Pero si un proceso cerebral no es interno, entonces qué es interno? Fijémonos en un ejemplo concreto. La investigación que se encarga del estudio del dolor no se distingue, en principio, de cualquier otra investigación empírica. Un científico podría llegar a saber que a una persona le duele una pierna mediante algún tipo de escáner neuronal; sin embargo, la relación que ambas personas mantienen con dicho estado mental es muy distinta. El científico ni siquiera tiene el dolor³² que identifica. Pero esto no genera *per se* ningún problema; A no puede experimentar el dolor de B, puesto que A no es B. Esto no es extraño; por el contrario, el falseamiento de este principio constituiría un gran misterio. Un mismo dolor no puede ser compartido por dos personas. Del mismo modo, A y B no pueden compartir la misma vida, aunque puedan tener una vida cualitativamente idéntica, y la historia de Polonia no puede ser la historia de España.

Se necesita otro punto de vista para poder formular una distinción filosóficamente interesante. Volviendo al ejemplo del médico, cabría preguntarse cómo los médicos pueden llegar a saber mediante un escáner neuronal que una persona tiene dolor en un momento dado. Remitirnos a los avances de la neurología no es una respuesta, ya que lo que tratamos de comprender es, precisamente, qué hace posible ese avance. ¿Cuál sería la situación del médico si dispusiera del escáner, pero no conociera cuál es el patrón a tener en cuenta para identificar un dolor? ¿Y cómo puede llegar a reconocer ese patrón?

³¹ “Intentar definir la experiencia consciente en términos de nociones más primitivas es infructífero. Del mismo modo, se podría intentar definir lo *material* o el *espacio* en términos de algo más fundamental. Lo mejor que podemos hacer es dar ilustraciones y caracterizaciones que se encuentren en el mismo nivel. Estas caracterizaciones no pueden calificar como verdaderas definiciones, dada su naturaleza circular implícita, pero pueden ayudar a determinar aquello de lo que se habla. Asumo que todo lector tiene experiencias conscientes de ella o él mismo. Si todo va bien, estas caracterizaciones ayudarán a establecer que son sólo *esas* de las que hablamos” (Chalmers 1996: 4).

³² Lo que aquí afirmamos no debe ser interpretado como un posicionamiento respecto al debate de la relación mente/cuerpo. Simplemente, trato de hacer inteligible la distinción de Chalmers.

Una respuesta de sentido común tendría en cuenta que disponemos de criterios muy claros para distinguir dolor en otros. Una forma habitual de saber cuándo a alguien le duele algo es prestar atención a lo que dice; este sería un buen método para identificar un patrón neuronal vinculado al dolor y, en general, a un estado mental. De hecho, hay algunas intervenciones en el cerebro que sólo se pueden llevar a cabo si el paciente está despierto y puede responder a las preguntas que se le van formulando. Otro criterio claro para identificar un dolor es la expresión no verbal, tanto en personas como en otros seres vivos. Un alarido, o el cuerpo de un animal retorciéndose suelen ser criterios inequívocos de dolor. En cambio, la ausencia de expresión fundamenta la confianza en el uso de anestesia durante una operación o en el consumo de alimentos vegetales como prácticas que no producen dolor en otros seres. Por supuesto, siempre hay casos dudosos. Se puede mentir, e incluso se puede fingir una expresión de dolor; en el caso de los animales, al menos de la mayoría, esto no es posible, ¿pero no es posible confundir una serie de espasmos con una expresión genuina de dolor? También hay patologías sorprendentes en las que una persona aparentemente en coma mantiene una plena conciencia de lo que sucede a su alrededor. Sin embargo, ¿en qué sentido es especial este margen para la incertidumbre? Al fin y al cabo, se trata de casos muy específicos. No somos infalibles al tratar de identificar la experiencia consciente de otros seres vivos, pero una larga historia de discusión en torno al escepticismo muestra que esta problemática no está específicamente vinculada a la experiencia consciente. Podemos plantear la duda escéptica respecto a si mañana saldrá el sol y responder a este tipo de dudas es una tarea compleja. Pero en tal caso, el problema que tratamos de plantear debería ser mucho más general.

No obstante, hasta ahora, solo he hablado de la relación que mantenemos con la experiencia consciente de otros seres. ¿Pero qué ocurre con la relación con nuestros propios estados mentales? Evidentemente, no sólo identificamos el dolor de otros seres, sino que sentimos nuestro propio dolor en primera persona, sin necesidad de recurrir a un artefacto o a los criterios de identificación que usamos para otros seres; además, lo sentimos de una manera específica, muy diferente a la que experimentamos con una caricia.

Volvamos al ejemplo de Nagel. Son muchas las cosas que sabemos de la experiencia perceptiva de los murciélagos; sabemos, por ejemplo, que experimentan el mundo a través de un complejo aparato de ecolocalización del que nosotros carecemos, y que su percepción visual es muy pobre. Sin embargo, no disponemos de un conocimiento de cómo se siente un murciélago. La dificultad no estriba en imaginarnos a nosotros mismos en la piel de un murciélago, especulando acerca de cuál sería nuestra experiencia si gradualmente nos fuéramos transformando en un murciélago; lo que nos está vedado es, precisamente, el conocimiento de cómo se siente un murciélago desde el punto de vista de un murciélago.³³ Las cualidades de nuestra experiencia consciente, en este sentido específico, se conocen como cualidades fenoménicas o *qualia* (*quale* en singular). Por otro lado, parece que tener experiencia consciente es un factor necesario para poder adscribir estados mentales conscientes a otras personas, de modo que todo el conocimiento del mundo que no tenga que ver con lo fenoménico no permite, por sí solo, determinar en qué casos hay experiencia consciente:

Esto es, la conciencia es *sorprendente*: si todo lo que supiéramos al respecto fueran los hechos de la física, incluyendo los hechos de la dinámica y del procesamiento de información en sistemas complejos, no habría una razón convincente para postular la existencia de la conciencia. Si no fuera por nuestra evidencia directa en primera persona, la hipótesis parecería injustificada; casi mística, quizá. Sin embargo, sabemos, de un modo directo, que hay experiencia consciente. La cuestión es: ¿cómo conciliamos la experiencia consciente con todo lo demás que sabemos? (Chalmers 1996: 5)

Que la conciencia sea sorprendente puede entenderse de dos formas. En primer lugar, el hecho de que haya conciencia en sí es algo imprevisible a partir del resto

³³ Esta es la misma intuición que sustenta el problema de las otras mentes. El problema de llegar a comprender cómo otros pueden tener estados mentales semejantes a los de uno, como señaló Wittgenstein, no consiste en la dificultad para concebir cómo podríamos tener esos estados mentales en el cuerpo de otro (podemos concebir que nuestros nervios se extienden hasta conectar con el sistema nervioso de otra persona). El problema es, más bien, que en principio no podemos entender qué es tener los estados mentales de otra persona, siendo otra persona.

de hechos del mundo; pero, aun dando por sentado que hay experiencia consciente, ¿por qué una experiencia consciente tiene el carácter específico que tiene y no otro? ¿Por qué al ver sangre tengo la experiencia visual de color que tengo y no la que tengo al observar el mar?

Chalmers parece estar acotando un fenómeno muy concreto y familiar, cuya comprensión se sustrae, a la vez, a un método de investigación estándar, ya que apunta a hechos que no están abiertos para ser inspeccionados. Esta limitación se relaciona con la idea de que la evidencia proporcionada por el punto de vista de la primera persona no puede reducirse al punto de vista de la tercera persona. Sin embargo, hay algo paradójico en el hecho de que en lo que trata de ser una caracterización inicial de la experiencia consciente encontremos incluida la formulación de un problema, a saber, nuestra incapacidad epistémica para derivar un subconjunto de los hechos del mundo a partir del resto de los hechos. De entrada, en la medida en que Chalmers dota de contenido el concepto de conciencia en el que va a centrar su estudio, parece que no hay otro modo de hacerlo, sino formulando dicho problema. Pero esto plantea la siguiente dificultad: la comprensión de un problema filosófico requiere una comprensión previa de los conceptos que van a entrar en juego, sin embargo, en este caso, los conceptos parece que solo pueden comprenderse a la luz de la formulación de un problema determinado. Surge, pues, una amenaza de circularidad.

No obstante, Chalmers señala que, respecto a la experiencia consciente, todo lo que se puede hacer es ofrecer ilustraciones y reconoce que es difícil dar definiciones de la misma que no incurran en circularidad. Una explicación de esto podría radicar en que la conciencia es un fenómeno tan básico y tan familiar que puede considerarse como primitivo. Sin embargo, el rasgo central de la conciencia, *what-it-is-likeness*, no parece que sea primitivo. Por otra parte, si el punto de partida ya se encuentra vinculado a la formulación de un problema metafísico, ¿en qué medida éste último tendrá una relación con los problemas que se planteen posteriormente? ¿Estarán los argumentos a favor de la irreductibilidad de la experiencia consciente viciados en su formulación al cometer una petición de principio? En cualquier caso, no se pueden extraer conclusiones tan

precipitadamente. Una formulación provisional de una noción puede ser muy general y no gozar de un sentido completo hasta que se aplica a casos particulares y se contrasta con formulaciones alternativas; en muchas ocasiones, un concepto solo cobra pleno sentido cuando es fructífero en la comprensión teórica. De momento, solo he presentado de un modo muy abstracto al planteamiento de Chalmers. Veamos con mayor detalle y fijándonos en diferentes instancias de cualidades fenoménicas en qué consiste el *explanandum* de Chalmers.

2.2. El catálogo y los dos conceptos de lo mental

Repasemos brevemente la serie de estados mentales que Chalmers considera como variedades de experiencia consciente y algunos de sus rasgos destacables:

Experiencias visuales: entre estas, las experiencias de color son paradigmáticas, aunque también hay que tener en cuenta las experiencias de forma, profundidad, brillo, etc.

Experiencias auditivas: se caracterizan, según Chalmers, por la peculiaridad de que, a diferencia de la experiencia visual, no parecen corresponderse con la estructura del mundo. Otro rasgo a destacar de estas experiencias es que, aunque se puedan descomponer los distintos sonidos de los que tenemos experiencia en tonos, notas, etc., en su conjunto pueden llegar a producir una experiencia cualitativa única, como en el caso de la música.

Sensaciones táctiles: el sentido del tacto proporciona una gran variedad de experiencias en función de las distintas texturas existentes; basta con pensar en el contraste entre el tacto del terciopelo y el de un metal frío.

Sensaciones olfativas: se caracterizan, según Chalmers, por carecer de estructura y por tener un carácter primitivo; de algún modo, esta carencia de estructura hace muy difícil describir sensaciones de olor.

Sensaciones del gusto: destaca como un sentido que engloba una gran variedad de experiencias, a pesar de que todas ellas parecen generarse a partir de un espacio de cuatro dimensiones: dulce, salado, amargo y agrio.

Otras sensaciones: las sensaciones a tener en cuenta además de las proporcionadas por los sentidos son innumerables; Chalmers destaca, no obstante, las sensaciones de frío y calor, las sensaciones corporales, como un picor, la sensación asociada a la necesidad de orinar, o el orgasmo. El dolor merece una mención aparte en la medida en que reúne unas características que lo hacen especialmente interesante como objeto de reflexión filosófica. Por una parte, es difícil hacerlo corresponder con una determinada estructura, en el mundo o en el cuerpo, aunque por otro lado se suele situar en una parte del cuerpo, lo que lo hace específicamente subjetivo.

Imágenes mentales: estas no solo incluyen aquellas asociadas a la imaginación, sino fenómenos como las postimágenes que surgen después de mirar algo brillante, o los patrones de colores que pueden aparecer al cerrar los ojos. Chalmers señala que lo específico de este tipo de experiencias es que son generadas “internamente”.

Emociones: las emociones no tienen un carácter tan local como las sensaciones, sino que más bien se caracterizan por permear y colorear nuestras experiencias. Por otra parte, parece que algunas emociones tienen ciertos aspectos que se asocian a los aspectos más cognitivos de la mente; pensemos, por ejemplo, en el miedo ante un peligro.³⁴

Otras experiencias: Chalmers tiene en cuenta dos tipos de experiencia a las que no se asocia tan fácilmente la conciencia, al menos en el sentido central esbozado más

³⁴ La dimensión cognitiva de las emociones ha sido subestimada a lo largo de la tradición de pensamiento cartesiano. En este sentido, el libro de Antonio Damasio *The Error of Descartes* (1994) no sólo trata de mostrar los aspectos cognitivos de las emociones, sino que éstas son esenciales para que los procesos cognitivos puedan funcionar.

arriba. El primero son los pensamientos conscientes, como las creencias. ¿Qué diferencia una creencia sobre un león de una creencia sobre una pera? Evidentemente el objeto, ¿pero podemos hablar de la existencia de *qualia* asociados a los pensamientos, algo así como cualidades leoniles frente a cualidades caninas de nuestras creencias? Y en tal caso, ¿hay diferencias entre dichas cualidades que resulten relevantes para la clasificación de los pensamientos? Dejaré de lado esta cuestión. El otro tipo de experiencia es el sentido del yo. Sin embargo, considerar que hay una experiencia del yo es algo muy controvertido. Chalmers habla tentativamente de “un tipo de zumbido de fondo”, aunque tiene un carácter intangible y es muy difícil de fijar. La reflexión filosófica sobre la conciencia del yo es un tema lo suficientemente complejo para merecer un estudio aparte, aunque más adelante diré algo al respecto.

Chalmers no pretende confeccionar una lista exhaustiva de todos los estados mentales abarcados por la experiencia consciente y reconoce que deja fuera muchos otros. En cualquier caso, el listado tiene un carácter más bien ilustrativo, ya que con vistas a la discusión posterior cualquiera de los ejemplos mencionados servirá igualmente a su propósito: “Al tratar los misterios filosóficos asociados con la experiencia consciente, una simple sensación de color plantea los problemas tan profundamente como la experiencia de uno al escuchar una coral de Bach” (Chalmers 1996: 11).

No obstante, dejando de lado los ejemplos más controvertidos de estados mentales a los que se les asocia *qualia*, como las creencias, no está muy claro que haya un aspecto bien definido que pueda dar unidad al catálogo de las experiencias conscientes. ¿Qué significa que tanto el olor a basura podrida como una gran alegría o la experiencia generada por la música de Bach coinciden en que ambas son de un modo para alguien (*they are like something*)? ¿En caso de que aceptáramos una clasificación tal podríamos considerarla como relevante para fijar un objeto digno de estudio? De entrada, podemos llevar a cabo clasificaciones extrañas, que incluyan bajo un mismo concepto una rueda de bicicleta, un cuadro y una nariz de payaso. Sin embargo, necesitamos comprender mejor cuál es el

aspecto subjetivo de nuestra experiencia que Chalmers considera el centro de su estudio.

De la mayor parte de ítems presentados en el catálogo hay numerosas y detalladas investigaciones científicas que podrían englobarse dentro del campo de estudio del trabajo interdisciplinar conocido como ciencia cognitiva. Pero afirmar que el objeto de estudio de las ciencias cognitivas coincide con lo mental no ayuda demasiado por dos razones: en primer lugar, puede haber distintas concepciones de lo mental, de modo que sostener que algo es mental no constituye, en sí mismo, una afirmación substantiva; en segundo lugar, tras la comprensión de ciertos fenómenos como la dimensión inconsciente de la mente, es evidente que la experiencia consciente, sin comprometernos con una noción específica de esta, no coincide con lo que intuitivamente consideramos como mental. Pensemos en los casos de percepción subliminal o en los pensamientos inconsciente. Chalmers señala que a lo largo de la historia de la filosofía ha habido una confusión entre dos conceptos de lo mental que debe ser evitada. Esta confusión no coincide, como veremos, con la identificación de lo mental y lo consciente, entendido en un sentido intuitivo:

En la base de esto se encuentran dos conceptos muy distintos de la mente. El primero es el *fenoménico*. Este es el concepto de la mente como experiencia consciente y de un estado mental como un estado mental experimentado conscientemente. Este es el aspecto más desconcertante de la mente y en el que me voy a concentrar, si bien no agota lo mental. El segundo es el concepto *psicológico* de la mente. Este es el concepto de la mente como la base causal o explicativa de la conducta. Un estado es mental, en este sentido, si juega el tipo adecuado de rol causal en la generación de conducta, o al menos juega un rol apropiado en la explicación de la conducta. Según el concepto psicológico, poco importa si el estado consciente tiene una cualidad consciente o no. Lo que importa es el rol que juega en una economía cognitiva.

En el concepto fenoménico, la mente se caracteriza por cómo se *siente*. En el concepto psicológico, la mente se caracteriza por lo que *hace*. No debería haber una cuestión de competencia entre estas dos nociones de la mente. Ninguna de

ellas es el análisis correcto de la mente. Las dos cubren fenómenos diferentes, y ambos son completamente reales. (Chalmers 1996: 11)

De entrada, esta distinción no añade nada nuevo a la noción de conciencia que estamos tratando de dilucidar, sino que reitera su rasgo central, a saber, que los estados mentales conscientes se sienten de una determinada manera, son de un modo específico para alguien, y esto es lo que Chalmers denomina el concepto fenoménico de lo mental. Sin embargo, la distinción determina el objeto de estudio de un modo negativo. Por una parte, desambigua notablemente a qué nos referimos cuando hablamos de la experiencia consciente; ya que, aunque desde hace tiempo se acepta ampliamente que lo mental no coincide con lo consciente, esta distinción todavía no precipita en un concepto claro de conciencia.

Es evidente que la experiencia consciente cumple una función vital fundamental que se manifiesta en nuestro comportamiento y en nuestra capacidad para llevar a cabo ciertas actividades, como conducir un vehículo, estar alerta o poder comprender lo que se nos dice; este es el concepto psicológico de conciencia. Pero, precisamente, este sentido de conciencia, que además coincide en gran parte con lo que ordinariamente entendemos por esta, no debe ser confundido con el concepto fenoménico de lo mental. Las cualidades fenoménicas de la experiencia no se caracterizan por regular la conducta. Cabría diferenciar distintas nociones de conciencia que se engloban en el aspecto psicológico de lo mental y que deben distinguirse del aspecto fenoménico. El aspecto psicológico incluiría acepciones de conciencia como la vigilia, la introspección, la habilidad para informar verbalmente sobre los propios estados mentales, la autoconciencia, la atención, el control voluntario de la acción o el conocimiento. De un modo general, Chalmers reserva el término *awareness* para referirse a la conciencia entendida en su acepción psicológica. Y esta distinción responde a la cuestión que planteábamos en referencia al objeto de estudio de las ciencias cognitivas. Estas últimas tratan de desentrañar cómo la mente constituye la causa interna del comportamiento y, en este sentido, no tendrían nada que decir respecto al concepto fenoménico de lo mental, sino que su objeto de estudio es el concepto psicológico de lo mental.

Chalmers asume la idea de que los conceptos psicológico y fenoménico agotan el ámbito de lo mental, en el sentido en que dos individuos idénticos en todo lo concerniente a dichos aspectos son idénticos mentalmente, aunque puntualiza que habría que añadir algunas propiedades relacionales en la medida en que el contenido de estados mentales como las creencias depende de ciertas propiedades del entorno del sujeto. Respecto al ámbito de lo psicológico, Chalmers se compromete con la idea de que el marco teórico del funcionalismo es correcto, aunque esto no implica que las ciencias cognitivas hayan terminado su trabajo. Todavía hay un gran número de problemas por resolver dentro de dicho marco, si bien ya se han dado los primeros pasos en la dirección adecuada para resolverlos, y no se trata de problemas metafísicos interesantes. Este es el conjunto de cuestiones que Chalmers llama problemas fáciles (*easy problems*). Aunque pueda tratarse de problemas muy difíciles de resolver, su dificultad es puramente técnica y no plantean ningún enigma filosófico. Este compromiso con el éxito de la empresa cognitiva así entendida resulta muy controvertido en la medida en que el funcionalismo, a pesar de su gran atractivo teórico³⁵, tiene que afrontar grandes problemas. No trataré esta cuestión aquí, aunque señalaré que, hasta el momento, no se han logrado dar definiciones funcionales completas de ningún estado mental y, por otra parte, hay dificultades para defender coherentemente la causalidad de lo mental en un marco funcionalista, poniéndose en peligro el fundamento del mismo proyecto cognitivo.

En cualquier caso, fuera cual fuera el resultado final de esta problemática, la cuestión que plantea Chalmers permanecería intacta. Por mucho que la empresa cognitiva, tal y como la presenté llegara a resolver y explicar todas las cuestiones

³⁵ Las definiciones funcionales de propiedades mentales ofrecen un gran atractivo teórico por tres razones básicas: en primer lugar, la flexibilidad del requisito funcional no impide considerar lo mental como algo interno ni implica, de entrada, una reducción de la mente a algo que no sea mental, permitiendo la causalidad entre estados mentales y evitando las objeciones básicas del conductismo; en segundo lugar, al definirse las propiedades funcionales en términos de un rol causal, el marco funcionalista no sólo es compatible con el marco fisicista dominante, sino que es especialmente adecuado para vincular propiedades físicas con propiedades mentales, de modo que las últimas dependan de las primeras; finalmente, las propiedades funcionales no se vinculan necesariamente a un único realizador, de manera que pueden afrontar el problema de la múltiple realizabilidad planteado a las teorías de identidad de tipos.

vinculadas al aspecto psicológico de lo mental, todavía resultaría incomprensible la existencia de la fenomenología, y éste es el *hard problem* de la conciencia: “El progreso impresionante de las ciencias físicas y cognitivas no ha arrojado luz de un modo significativo respecto a la cuestión de cómo y por qué el funcionamiento cognitivo es acompañado por experiencia consciente” (Chalmers 1996: 22).

2.3. La doble vida de los estados mentales y la co-ocurrencia de las propiedades psicológicas y fenoménicas

Si volvemos a la serie de ejemplos que hemos tenido en cuenta para ilustrar a qué nos referimos cuando hablamos de las cualidades fenoménicas de la experiencia, surge un problema al aplicar los dos conceptos de lo mental. El aspecto psicológico del dolor, tal y como lo entiende Chalmers, podría definirse provisionalmente como un estado producido por el daño en el tejido biológico que causa un gran malestar y da lugar a la retirada inmediata de la parte del cuerpo sometida al daño; la neurología nos puede enseñar, a su vez, que dicho rol es desempeñado, al menos en nuestra especie, por un determinado tipo de fibras nerviosas. Sin embargo, parece que esto no es todo lo que se puede decir respecto a un dolor; cuando hablamos de dolor nos referimos a la sensación específica que distinguimos de un picor o de una sensación de frío, y ése es su aspecto fenoménico. Incluso parece que, si el dolor es lo suficientemente intenso, éste se impone de tal modo que prácticamente no podemos atender a otra cosa. Ahora bien, ¿en qué medida puede decirse que el dolor tiene tanto un aspecto fenoménico como psicológico? Y si es así, ¿tienen la misma importancia ambos aspectos para determinar si un estado es dolor? Imaginemos que hay un tipo de anestesia que elimina por completo la sensación de dolor pero que tiene el curioso efecto de producir un espasmo semejante al movimiento que llevamos a cabo cuando algo nos duele y, además, ese espasmo se produce en aquellos casos en los que se sufre un daño corporal por parte del sujeto, ¿podemos hablar en tal caso de dolor? ¿Y qué ocurre con aquellos casos en los que hay una sensación de dolor, pero no viene acompañada de las disposiciones conductuales habituales? Pensemos en los casos de dolores que duran todo un día y de los que dejamos de ser conscientes debido a su

duración; también se sabe que algunas drogas producen el efecto curioso de mantener una sensación de dolor que ya no resulta molesta. ¿Tiene sentido seguir hablando en estos casos de dolor? Chalmers lleva a cabo la siguiente puntualización al respecto:

Nada importante descansa en la decisión semántica respecto a si alguna cualidad fenoménica es *realmente* esencial para que algo cuente como dolor. En lugar de esto, podemos reconocer los diferentes componentes asociados con un concepto y distinguirlos explícitamente, hablando, por ejemplo, de “dolor fenoménico” y “dolor psicológico”. Nuestro concepto cotidiano de dolor supuestamente combina ambos en una sutil combinación que ha sido ponderada, pero para la discusión filosófica las cosas son mucho más claras si las mantenemos separadas. (Chalmers 1996: 17)

De acuerdo con esta aclaración, no hay ninguna cuestión importante implicada respecto a qué considerar como instancia de un determinado estado mental en relación a los dos conceptos de lo mental, siempre y cuando se tenga en cuenta a qué componente de lo mental nos referimos, en cada caso, cuando hacemos filosofía. Por otra parte, puede decirse que algunos estados mentales tienen una mayor tendencia a identificarse con uno de los dos componentes. En el caso de la sensación, parece que el componente fenoménico tiene mayor peso; en el caso de las disposiciones psicológicas o las habilidades mentales se trata, más bien, de una noción psicológica. Las actitudes proposicionales parecen, en principio, estar más vinculadas al aspecto psicológico de lo mental, aunque plantean una problemática más compleja que no trataré aquí.³⁶

Es un hecho que en la práctica cotidiana no hacemos una distinción entre los dos conceptos de lo mental. Igualmente, en el debate filosófico ha habido una

³⁶ Aunque, según la terminología de Chalmers, nuestras creencias estén siempre acompañadas de cualidades fenoménicas específicas, no parece que tengamos una creencia u otra en virtud de una diferencia en la fenomenología. No obstante, hay una cuestión más general que tiene que ver con la posibilidad de que un ser carente de fenomenología tenga estados intencionales. Algunos filósofos, como John Searle, niegan que esto sea posible, de manera que en los casos en los que esto parece suceder, como pueda ser en el caso de un ordenador, todo lo que hay es una intencionalidad “como si” (as if).

tendencia a confundir ambos conceptos, de modo que la distinción no es evidente incluso después de haber iniciado una reflexión teórica. Chalmers es consciente de esta cuestión, aunque no considera que esto suponga una dificultad relevante para su planteamiento. Su explicación es la siguiente:

Es un hecho respecto a la mente humana que siempre que se instancia alguna propiedad fenoménica, alguna propiedad psicológica correspondiente se instancia. La experiencia consciente no ocurre en el vacío. Está siempre ligada al procesamiento cognitivo y es probable que, en algún sentido, surja de dicho procesamiento. Siempre que uno tiene una sensación, por ejemplo, algún procesamiento de información tiene lugar: una percepción correspondiente, si así se quiere. De un modo similar, cuando uno tiene la experiencia consciente de felicidad, el rol funcional asociado con la felicidad es desempeñado, generalmente, por un estado interno. Quizá sea lógicamente posible que uno pueda tener la experiencia sin la causalidad, pero parece ser un hecho empírico que vayan juntas.

Dada esta co-ocurrencia, los temerosos podrían estar tentados de cuestionar si se está haciendo una distinción real. Pero está claro que aquí hay, al menos, una distinción conceptual, incluso si las extensiones de los conceptos involucrados parecen ir juntas. Uno puede preguntarse cómo explicar la cualidad fenoménica y preguntarse cómo explicar el desempeño del rol causal, y estas son dos preguntas distintas. (Chalmers 1996: 22)

Lo que explica la confusión de los dos conceptos de lo mental es, entonces, el hecho empírico de que haya una co-ocurrencia de las instancias de cada uno de los conceptos de lo mental. Esta explicación parece responder a las cuestiones que planteábamos más arriba: hay un rasgo del mundo, en relación a nuestros estados mentales, que se refleja en nuestra manera de conceptualizarlo. Dado el vínculo que hay entre lo psicológico y lo fenoménico, resulta plausible que no sea necesario, en las situaciones de la práctica cotidiana, trazar una distinción como la descrita. En este sentido, podemos encontrar fenómenos análogos en el uso de nuestros conceptos que no sólo no constituyen casos aislados, sino que permean nuestras prácticas. Pensemos por ejemplo en el modo que tenemos de referirnos a las alergias. Es bastante habitual oír hablar de la alergia al polvo; sin embargo, esta

expresión es, como mínimo, inexacta. Lo que provoca la alergia realmente son los ácaros del polvo. Es un hecho empírico que la presencia de polvo va acompañada de ácaros, de modo que teniendo en cuenta que el interés práctico en las alergias se vincula principalmente con la evitación de las molestias que ésta provoca, y dado que evitando la presencia de polvo se evita la presencia de ácaros, no es extraño que la alergia se identifique como alergia al polvo, ya que el vínculo que resulta relevante para la prevención de la alergia es el que hay entre el polvo y los síntomas de la alergia. Parece que, en general, hay un espacio para redefinir nuestros conceptos a la luz de intereses pertenecientes a distintas prácticas, pero esto no parece plantear ningún problema profundo. Además, Chalmers puede explicar de este modo la idea de que la experiencia consciente no flota libremente, sino que tenemos la fuerte intuición de que se encuentra robustamente vinculada a los procesos cerebrales y al cuerpo. El tipo de nexo es empírico y, por tanto, contingente; pero a falta de un argumento esto no supone ningún problema.

El principio de la co-ocurrencia refleja, según Chalmers, un hecho notable de nuestra experiencia consciente, recogiendo una idea a la que ya apuntó Gilbert Ryle: no disponemos de un lenguaje específico para describir cualidades fenoménicas, de modo que la referencia de éstas se fija mediante propiedades externas o roles causales. Chalmers señala que ésta podría ser una interpretación de la afirmación wittgensteiniana. Ahora bien, Chalmers continúa, este hecho no debería llevarnos a suponer que esos criterios que nos permiten fijar la referencia de lo fenoménico es todo lo que hay y trata de ilustrar la implausibilidad de esta tesis con el siguiente ejemplo: el emblema de un Rolls-Royce es el tipo de emblema que normalmente se encuentra en un coche Rolls-Royce, pero esto no implica que dicho emblema sea el mismo coche. ¿Pero sirve este ejemplo para comprender la plausibilidad del principio de co-ocurrencia de lo psicológico y lo fenoménico? A continuación, plantearé una serie de problemas que dificultarán la formulación de la distinción entre los dos conceptos de lo mental.

2.4. El problema de la individuación de las cualidades fenoménicas

Un rasgo especial del aspecto fenoménico de lo mental es que, para poder hablar de cualidades fenoménicas, nos servimos de ejemplos de lo psicológico. Esta es una asimetría notable respecto a los ejemplos propuestos y se trata de un rasgo central de la experiencia consciente.

En el caso del emblema de Rolls-Royce, es evidente que la identidad de los coches no se fija, exclusivamente, por la referencia al emblema; además, la referencia del emblema también se fija independientemente de que esté situado en el coche. Un emblema puede desprenderse y se puede seguir identificando perfectamente. En el caso del ejemplo de las alergias es evidente que los ácaros pueden identificarse sin necesidad de apelar a la existencia de polvo; quizá, incluso es posible modificarlos genéticamente para que ya no se encuentren en el polvo.

Ahora bien, ¿qué sucede en el caso de las cualidades fenoménicas? ¿Puede fijarse su identidad sin referencia a un criterio externo? ¿Y es importante que esto sea posible para trazar la distinción de Chalmers? En relación a esta última pregunta cabría decir que la respuesta es afirmativa en la medida en que Chalmers considera que la relación de lo fenoménico con lo psicológico es una relación empírica y, por ello, contingente, como la que hay entre un Rolls-Royce y su emblema. Si esto es así, es necesario poder fijar la identidad de las cualidades fenoménicas para después poder investigar el tipo de relaciones contingentes que estas pueden mantener con lo demás. Este requisito no debe confundirse con una descripción del mundo caracterizada por la existencia de una serie de objetos que se identifican con independencia de todo lo demás; el requisito es mucho más laxo. Todo lo que me limito a afirmar es que para poder reconocer una relación contingente R entre A y B , las condiciones de identidad de A no pueden depender de la existencia de la relación R . En tal caso, dicha relación ya no es contingente sino lógicamente necesaria.

En el caso del emblema, es evidente que ambos objetos tienen condiciones de identidad independientes del hecho de que uno de ellos sea un Rolls-Royce y el otro un emblema de Rolls-Royce. Como ya he dicho, un emblema es un objeto que

se puede separar del coche y sigue conservando su identidad.³⁷ Ahora bien, en el caso de la relación entre cualidades fenoménicas y procesos psicológicos, ¿es posible determinar criterios de identidad de ambos que sean independientes de la relación empírica de co-ocurrencia? Si la respuesta a esta pregunta es negativa, como trataré de mostrar a continuación, la distinción de Chalmers no es coherente.

De entrada, parece que los procesos psicológicos de la mente tienen unas condiciones de identidad independientes. Pensemos, de nuevo, en la definición funcional que dimos del dolor como el estado producido por un daño en los tejidos corporales que causa la retirada de la parte del cuerpo dañada. ¿Podemos decir que el patrón así descrito posee condiciones independientes de identidad? ¿Qué cuenta como un movimiento de retirada al sufrir un daño? Pensemos en el caso de la anestesia que produce un espasmo semejante a un gesto de dolor, ¿ejemplifica dicho espasmo el patrón adecuado? ¿Diríamos que el movimiento de un brazo mecánico al recibir un golpe cuenta como el tipo adecuado de movimiento? Hay plantas que al contacto con un objeto se retraen, ¿qué nos impide decir que esto es una ejemplificación, al menos psicológica, de dolor? Por supuesto, todos estos movimientos pueden identificarse bajo distintas descripciones, del mismo modo que el emblema de Rolls-Royce es un trozo de metal entre otros, o una pequeña figura humana; pero lo importante es que puedan identificarse como el tipo de estado que se incluye en el concepto de dolor.

La pregunta no es, pues, si se pueden describir patrones de estados que coinciden con los mismos particulares que consideramos mentales, sino en virtud de qué se eligen unos patrones y no otros. Incluso si no tenemos en cuenta casos extraños, cualquier comportamiento puede generar este problema: ¿cuándo comienza y cuándo acaba una conducta de dolor? Dada la ocurrencia simultánea de dos estados mentales M y S en un tiempo incluido entre t_1 y t_2 , ¿qué partes del

³⁷ Incluso sus condiciones de identidad son independientes del hecho de que haya coches Rolls-Royce. Imaginemos un mundo posible en el que la empresa Rolls-Royce, antes de fabricar todos los coches de su primer modelo, fabrica los emblemas. En ese momento, los emblemas, considerados como objetos, tienen condiciones de identidad independientes de la existencia de coches. Por supuesto, su identidad no es independiente del hecho de que haya una empresa dispuesta a fabricar coches en los que estos objetos figurarán como emblemas, pero este tipo de dependencia no afecta a la relación contingente entre los coches y los emblemas.

movimiento de mi brazo que se produce entre el momento t_1 y t_2 forman parte del componente psicológico del estado mental M y no del estado mental S, y cuáles son meros espasmos? ¿Y qué quiere decir que se produce una conducta de dolor o una conducta de tristeza sin la sensación de dolor o sin el sentimiento de tristeza?

Parece que Chalmers, cuando habla de la co-ocurrencia, está asumiendo una relación de dependencia unidireccional; de hecho, el ejemplo que propone para ilustrar dicha co-ocurrencia es el de experiencias que empíricamente siempre van acompañadas del correspondiente procesamiento cognitivo. Pero si tratáramos de partir de algún tipo de procesamiento cognitivo, ¿cómo podríamos estar seguros de que hay una experiencia fenoménica acompañando a dicho procesamiento? ¿Cómo podríamos responder satisfactoriamente a las preguntas del párrafo anterior? En los ejemplos analizados, inadvertidamente, se está otorgando a la fenomenología un papel dominante en la individuación de estados mentales. Es posible que, como un hecho empírico respecto a nuestro modo de formar algunos conceptos, sea necesario que tengamos también contacto con el aspecto fenoménico de lo mental, y esto, a su vez, nos permita reconocer patrones psicológicos que antes no distinguíamos. Y es posible que, una vez adquirimos esta capacidad, ya no necesitemos recurrir a la fenomenología para identificar dichos patrones.³⁸

La eficacia de una respuesta de este tipo es dudosa, ya que no explicaría por qué nuestra forma efectiva de formar conceptos se encuentra constreñida de esta forma, ni en qué medida este tipo de restricción no es lógico. En cualquier caso, para poder articular esta respuesta es necesario que las instancias particulares de las cualidades fenoménicas puedan individuarse independientemente del concepto psicológico de lo mental ya que, de otro modo, ¿cómo podríamos llegar a identificar estados mentales si lógicamente la fenomenología es anterior? Esto nos lleva al otro lado de la co-ocurrencia: ¿cuáles son las condiciones de identidad de las

³⁸ Chalmers (2000) trata de dar con una definición de correlato de la conciencia que apunta a un sistema que no necesariamente se correlaciona exclusivamente con la experiencia consciente y tampoco tiene por qué tratarse de un área específica del cerebro. Podría objetarse que, en la discusión de este capítulo estoy simplificando demasiado el tipo de candidato a correlato de la conciencia que la ciencia aspira a encontrar. Sin embargo, en la argumentación presentada, esta complejidad no desempeña ningún rol, y el argumento podría repetirse igualmente, en la medida en que necesitamos un criterio para poder afirmar que determinado proceso o conjunto de procesos en el cerebro constituyen una unidad que puede correlacionarse con la experiencia consciente.

instancias de cualidades fenoménicas? Por definición y, teniendo en cuenta lo dicho, ha quedado excluida cualquier referencia a criterios externos. Esta limitación sólo apunta, según Chalmers, a nuestra incapacidad para referirnos verbalmente a cualidades fenoménicas, pero, entonces, ¿cómo somos capaces de diferenciar experiencias y tipos de experiencias sin atender a criterios externos?

Decir que las cualidades fenoménicas poseen un carácter intrínseco no nos ayuda a entender el mecanismo epistémico del que nos servimos para poder clasificar experiencias y, al menos de momento, no disponemos de un modelo paralelo que nos permita entender en qué consiste ese carácter intrínseco. Chalmers señala que el conocimiento que tenemos de las cualidades fenoménicas de nuestra experiencia es más directo que el que tenemos respecto a cualquier otra cosa, y esto es lo que explica que la experiencia consciente sea una sorpresa desde el punto de vista de la tercera persona; es necesaria la evidencia adicional proporcionada desde el punto de vista de la primera persona. Quizá la forma de relacionarnos con nuestra propia experiencia consciente es tan primitiva que no es posible comprenderla mejor. Sin embargo, recurrir a la evidencia de la primera persona de este modo particular introduce una dificultad lógica todavía mayor, como mostraré a continuación.³⁹

Consideremos, en este sentido, el análisis de las condiciones de posibilidad de nuestras atribuciones de estados mentales que nos ofrece Peter F. Strawson (1959). Esta investigación es una parte de la descripción nuestro esquema conceptual general. Según Strawson, los cuerpos físicos y las personas son los particulares básicos de nuestro esquema conceptual, en la medida en que los otros particulares solo pueden ser concebidos por referencia a dichos particulares básicos. Por ejemplo, para poder identificar un lugar determinado es necesario poder remitirnos a ciertos objetos físicos. En el caso de las personas, según Strawson, no podríamos referirnos a estados mentales sin la posibilidad de atribuirlos a personas. Esta es una idea bastante intuitiva, ya que los estados mentales no flotan en el vacío, como incluso reconoce Chalmers. Parece que la idea

³⁹ El argumento que a continuación presentamos es una aplicación del argumento que P. F. Strawson utiliza en Strawson (1959) a favor de la comprensión del concepto de “persona” como un concepto más básico que los de cuerpo y mente.

de un dolor desencarnado, por ejemplo, no tiene sentido. Como veremos, independientemente de la conclusión de que el concepto de persona sea un concepto básico, la argumentación de Strawson plantea un problema general que afecta a cierta forma de entender la mente y que hace ininteligible la doble concepción de lo mental de Chalmers.

Es evidente que el cuerpo de cada uno juega un rol crucial en la identificación de los estados mentales propios: no sentimos dolor cuando cualquier cuerpo recibe un golpe, sino que esto solo ocurre cuando un cuerpo C determinado es golpeado. Una forma de explicar por qué consideramos C como nuestro cuerpo es que este ocupa un rol causal especial en relación al mundo y nuestras experiencias conscientes. Ahora bien, hay dos preguntas que esta explicación no puede responder. En primer lugar, ¿por qué nuestros estados conscientes tienen que ser atribuidos a algo? En segundo lugar, ¿por qué atribuimos nuestros estados conscientes a las mismas cosas a las que atribuimos características físicas?

Una reacción a la idea de que atribuimos ambos tipos de estado a la misma cosa es que esta idea es el resultado de una confusión. Strawson hace referencia a dos posiciones que ejemplifican esta reacción. El primer tipo de respuesta, que puede atribuirse a Descartes, afirma que, en realidad, las características físicas y los estados mentales se atribuyen a sustancias distintas, de modo que la segunda pregunta está equivocada. De acuerdo al segundo tipo de concepción, incluso la primera pregunta estaría errada. Esta segunda concepción, que pudo haber sido defendida por Moritz Schlick y quizá Wittgenstein en algún momento, afirma que los estados mentales ni siquiera son atribuidos a un sujeto. El hecho de que el cuerpo de cada uno goce de un rol causal especial da lugar a la ilusión de posesión de estados mentales por parte de un sujeto, pero este sentido de posesión, que se reduce a una relación causal, genera la idea errónea de que los estados mentales se atribuyen a un sujeto E que no es físico, de manera que la relación entre E y los estados mentales no sería contingente sino necesaria. Podemos llamar a esta posición la concepción “sin sujeto”, según la cual, las dos preguntas planteadas en el párrafo anterior no tienen sentido. Sin embargo, Strawson trata de mostrar que

esta última posición no puede enunciarse coherentemente, examinemos su argumento.

Veíamos que la experiencia de cada persona se encuentra vinculada causalmente de un modo especial con un cuerpo determinado. Puesto que el vínculo entre dicho cuerpo y la experiencia de la persona es causal, es también contingente. Desde la posición “sin sujeto”, el vínculo entre las experiencias de una persona y su cuerpo se podría describir del siguiente modo: “Todas *mis* experiencias son tenidas por el cuerpo C” y por “tenidas” hay que entender que se vinculan causalmente y de un modo singular con el cuerpo C. Así formulado, el principio no dice lo que el teórico de la concepción “sin sujeto” pretende decir, ya que el posesivo *mis* no puede significar, simplemente, las experiencias que son tenidas por el cuerpo C. Esto daría lugar a la siguiente tautología: “Todas las experiencias que son tenidas por el cuerpo C son tenidas por el cuerpo C”. Pero el principio requerido tiene que ser contingente, no puede ser una tautología. Tampoco serviría decir “Todas las experiencias son tenidas por el cuerpo C”, ya que sería un principio claramente falso. Dicho de otro modo, el problema con el que se encuentra el teórico de esta posición es que no es posible ser consciente de las propias experiencias, como propias, si no somos capaces de entender, previamente, que nuestras experiencias no podían haber sido de otros. Esto implica que entendemos que sujetos distintos a nosotros mantengan el mismo tipo de relación con sus propios estados mentales o, al menos, ha de resultar inteligible la idea de que nuestras experiencias en primera persona no sean las únicas que existen, ya que, ¿qué puede querer decir que *mis* experiencias son *mías* si no entiendo que otros también tienen *sus* propias experiencias? ¿Cómo puedo entender la experiencia de los demás tomando como punto de partida *mi propio caso* si no entiendo cuál es *mi caso*?:

¿Ya que qué significa, ahora, “nuestro propio caso”? La idea de atribuirse a uno mismo, o en general, estados conscientes no tiene sentido, a menos que quien los atribuya sepa ya cómo atribuir, al menos, algunos estados conscientes a otros. De este modo, no puede argumentar extrayendo conclusiones en general “a partir de su propio caso” respecto a cómo hacer esto; puesto que, a menos que ya sepa cómo

hacer esto, no tiene una concepción de *su propio caso*, o de cualquier caso, esto es, de cualquier sujeto de experiencias. (Strawson 1959: 106)

Lo que plantea el argumento no es que una generalización a partir del caso propio sea inadecuada. El problema es que ni siquiera podríamos entender cuál es nuestro propio caso, ya que no entenderíamos de qué concepto es instancia nuestro propio caso. Por otra parte, es un hecho que cuando atribuimos estados mentales, sea a otros o a nosotros mismos, no nos servimos de un único criterio de atribución e incluso, cuando nos autoadscribimos un estado mental consciente, parece que ni siquiera nos servimos de un criterio que preceda a la atribución, simplemente lo hacemos. La consecuencia del argumento de Strawson es, en primer lugar, que las atribuciones de estados mentales significan lo mismo cuando se hacen a otros y cuando se trata de autoatribuciones. La segunda consecuencia es que los criterios que utilizamos para atribuir estados mentales a otras personas constituyen parte de su significado. La conclusión de Strawson no es que los estados mentales sean reducibles a estos criterios, sino que, si entendemos su significado, entendemos que son atribuibles tanto teniendo en cuenta criterios públicos como desde el punto de vista de la primera persona.

Siguiendo la argumentación de Strawson, el problema que plantea la distinción de Chalmers entre los dos conceptos de lo mental es que no puede describirse la co-ocurrencia entre los aspectos psicológico y fenoménico de lo mental como una relación contingente, ya que los criterios externos de atribución de conciencia no son independientes, conceptualmente, de la instanciación de experiencia consciente. Por supuesto, podría dejar de tomarse como punto de partida un conjunto de experiencias cuya identidad se considere independiente lógicamente de criterios públicos de atribución de las mismas. Lo que se negaría no es que éstas existan, sino su prioridad lógica, de modo que comprender que uno mantiene una determinada relación con *sus propias* experiencias implica ya, necesariamente, la capacidad de atribuir experiencias mentales a otros, y esto ineludiblemente supone el reconocimiento de un vínculo necesario entre las condiciones de identidad de nuestros estados mentales y ciertos criterios públicos.

Pero la aceptación de esta solución violaría el principio de co-ocurrencia, en la medida en que la co-ocurrencia es una relación contingente, y mostraría que la distinción de Chalmers no se puede trazar coherentemente.

En este capítulo, he tratado de comprender un concepto de experiencia consciente que en los debates filosóficos actuales plantea un desafío a nuestra comprensión general de las relaciones entre la mente y el resto del mundo; esta noción de experiencia se vincula al aspecto subjetivo de ésta, en el sentido más específico de *what-it-is-likeness*. Esta noción, a la que David Chalmers denomina el concepto fenoménico de lo mental, debe distinguirse del concepto psicológico de lo mental, el cual se identifica con el acceso de un sistema a la información con el fin de controlar su comportamiento y con el rol causal que determinado estado desempeña en la economía cognitiva de dicho sistema. Chalmers explica la tendencia a confundir ambos conceptos por su co-ocurrencia, de modo que la ejemplificación de cualidades fenoménicas viene acompañada, contingentemente, por procesos cognitivos. Por otra parte, la co-ocurrencia de los dos conceptos de lo mental, implica que las instancias de ambos conceptos tienen condiciones de identidad independientes entre sí; sin embargo, si la argumentación presentada es correcta, entonces no es posible individuar cualidades fenoménicas con independencia de criterios externos y, por tanto, no se puede formular coherentemente la distinción de Chalmers, ya que la identidad de nuestras experiencias conscientes depende de poder atribuir determinados tipos de experiencias a otras personas.

3. Argumentos a favor de la irreductibilidad de la conciencia

Hasta ahora he identificado el marco general que se encuentra implícito tanto en la formulación del problema de la conciencia como en toda visión que considere dicho problema inteligible; sin embargo, todavía no he presentado un argumento explícito que establezca el problema de la conciencia. Ya vimos que la caracterización del tipo de conciencia que Chalmers considera filosóficamente interesante y problemático, el aspecto fenoménico de la experiencia consciente, no es independiente de la formulación del mismo problema de la conciencia. A continuación, presentaré los argumentos principales de los que se sirve Chalmers para concluir que la experiencia fenoménica es irreducible a lo físico. De un modo general, y bajo la interpretación de Chalmers, estos argumentos apuntan conjuntamente a la conclusión de que los hechos relativos a la experiencia fenoménica no supervienen sobre los hechos físicos.

No obstante, hay diferencias entre los distintos argumentos por dos razones. En primer lugar, aunque todos los argumentos permitieran concluir que los hechos relativos a la experiencia consciente no supervienen sobre los hechos físicos, podrían establecer esta conclusión con matices distintos. Por ejemplo, algunos argumentos, concluyen que, dado cualquier conjunto de hechos físicos, no hay una necesidad lógica de que haya conciencia, mientras que otros argumentos apuntarían a la conclusión más débil de que, aunque un conjunto de hechos físicos sea suficiente para la existencia de experiencia fenoménica, no lo es para la determinación del carácter cualitativo específico de dicha experiencia. La segunda razón, más importante, es que los distintos argumentos se apoyan en consideraciones distintas. Algunos apelan a la concebibilidad como una vía que permite extraer conclusiones metafísicas, mientras que otros apelan a intuiciones epistemológicas o relacionadas con el análisis de los conceptos de estados mentales. Finalmente, los elementos fundamentales para poder expresar claramente un argumento antimaterialista serán la concebibilidad y su relación con la posibilidad metafísica. Con el fin de poder desarrollar dicho argumento, presentaré el marco semántico del que Chalmers se sirve para construir el

argumento y concluiré que la discusión filosófica relevante sobre la experiencia consciente se encuentra al margen de la validez de dicho marco semántico.

3.1 Los argumentos

Siguiendo a Chalmers, voy a exponer los distintos argumentos atendiendo al tipo de consideraciones en las que se fundamentan. Este criterio de clasificación, aunque útil para comenzar un estudio de los argumentos, no debe ser tomado como último, en la medida en que estos se encuentran interconectados y recurren a concepciones previas más básicas. Como mostraré, haber pasado por alto este hecho ha dado lugar a una serie de debates cada vez más complejos y especializados que, no obstante, han dejado intacto un punto de partida muy cargado metafísicamente. En primer lugar, presentaré dos argumentos que apelan a la concebibilidad de determinadas situaciones, como un test que permite extraer conclusiones respecto a la posibilidad lógica de éstas. Estos argumentos son los más conocidos y tratados en la bibliografía y, en parte, me he servido de ellos implícitamente para dotar de contenido al proyecto fisicista.

No está claro que los argumentos que voy a presentar sean equivalentes entre sí, aunque todos van a remitir a la intuición de que los hechos respecto a la experiencia consciente no se encuentran fijados por los hechos físicos o, según la terminología presentada, la intuición principal es que los hechos fenoménicos no supervienen sobre los hechos físicos.

3.1.1. La posibilidad lógica de los zombis

El primero de los argumentos es la posibilidad lógica de los zombis. Un zombi es un ser física y funcionalmente idéntico a nosotros que, no obstante, carece de experiencia fenoménica. Un zombi filosófico no tiene nada que ver con los muertos vivientes cinematográficos; estos últimos muestran un alto grado de disfuncionalidad en sus capacidades mentales y el control de su comportamiento que los hace fácilmente identificables. Por el contrario, un zombi es completamente

indistinguible de un ser humano normal, física, biológica y psicológicamente⁴⁰. Este tipo de zombi responderá adecuada e inteligentemente a cualquier pregunta que se le formule, reirá después de escuchar un chiste, dormirá, tendrá su propio carácter e incluso podría discutir sobre la experiencia fenoménica y sobre la posibilidad de los zombis. Ahora bien, aunque afirme que posee experiencia fenoménica (lo cual hará puesto que es indistinguible de nosotros) carecerá de esta; cuando se encuentre masticando regaliz, no se encontrará asociada la cualidad fenoménica del sabor del regaliz, y aunque dicho ser discrimine colores adecuadamente, lo hará sin que se le haga presente el carácter cualitativo específico de cada color. Dicho ser carecerá de lo que he denominado *what-it-is-likeness*⁴¹. Ahora bien, y este es el argumento, si podemos concebir dichos seres, entonces son posibles lógicamente, y por lo tanto el materialismo⁴² es falso, de manera que el argumento tendría la siguiente forma:

- (1) Los zombis son concebibles.
- (2) Si los zombis son concebibles, entonces son lógicamente posibles.
- (3) Si los zombis son lógicamente posibles, el materialismo es falso.
- (4) Por lo tanto, el materialismo es falso.

Aunque (3) no sería aceptada por muchos materialistas que consideran el materialismo como una verdad contingente, o válida solo en nuestro mundo, dado

⁴⁰ Para no crear malentendidos, y si no se indica otra cosa, cuando hablemos de lo psicológico nos referiremos al sentido definido por Chalmers, según el cual psicológico es todo aquello que tiene que ver con el uso de información con el fin de regular la conducta.

⁴¹ Ahora bien, esta carencia no debe confundirse con la carencia de una perspectiva. Un zombi, por ocupar un lugar concreto y situarse en una historia causal determinada, (y admitimos que puede tener estados intencionales sin fenomenología), en cierto sentido gozará de una perspectiva, aunque también esta idea es controvertida.

⁴² Aunque no son exactamente equivalentes, en lo sucesivo hablaremos indistintamente de materialismo y fisicismo. El materialismo alude directamente a la materia, de manera que podría considerarse, estrictamente, un fisicismo comprometido con una física en la que la materia es la noción fundamental; por el contrario, el fisicismo no se pronunciaría, de entrada, respecto a ninguna noción fundamental de la física.

el marco desarrollado en el capítulo anterior, podemos considerarla como una estipulación. Si los hechos relativos a la experiencia fenoménica no supervienen lógicamente sobre los hechos físicos, entonces el materialismo es falso. (3) podría figurar como conclusión del siguiente silogismo hipotético:

(1') Si los zombis son lógicamente posibles, entonces los hechos relativos a la experiencia fenoménica no supervienen lógicamente sobre los hechos físicos.

(2') Si los hechos relativos a la experiencia fenoménica no supervienen lógicamente sobre los hechos físicos, entonces el materialismo es falso.

(3) Si los zombis son lógicamente posibles, entonces el materialismo es falso.

Ya vimos que la tesis fisicista no goza de un contenido suficientemente preciso, en la medida en que la caracterización de lo que se considera físico no juega un papel determinante en la discusión. No obstante, aun si no conseguimos determinar una tesis materialista bien definida, parece que para cualquier noción relevante de "físico", en caso de que el argumento de la concebibilidad funcionara, se seguirían algunas consecuencias claras que afectarían a nuestras creencias de sentido común. Por ejemplo, se establecería que nuestras creencias respecto a la experiencia consciente de los demás podrían ser sistemáticamente erróneas.⁴³ La experiencia consciente dejaría de estar anclada al mundo de manera que, una vez aceptado este movimiento, no sólo no podría establecerse con garantías la existencia de conciencia en casos que podríamos considerar paradigmáticos, sino que sería difícil descartar la posibilidad de que exista conciencia en casos en los

⁴³ Una consecuencia práctica de este argumento que trataremos más adelante tiene que ver con el valor que asociamos a la experiencia consciente. Consideramos que la vida es digna de ser vivida en la medida en que hay cierto grado de conciencia y, del mismo modo, consideramos que un ser merece un trato determinado en virtud de su grado de conciencia; de manera que un ser carente de experiencia consciente sería semejante a un objeto. Ahora bien, en tal caso, una consecuencia del argumento de la concebibilidad sería que seres aparentemente semejantes a nosotros en todo podrían merecer el mismo trato que un objeto.

que en principio nos resulta claro que ésta no se da, como en el caso de una piedra que es golpeada.⁴⁴

La premisa (1), según Chalmers, no puede fundamentarse en un razonamiento más básico, sino que más bien, el peso de la prueba debería recaer en quienes consideren que los zombis no son concebibles. Según Chalmers, alguien que defienda la descripción coherente de una situación contrafáctica no puede hacer mucho más para argumentar a favor de su posibilidad. En todo caso, puede presentar algunos casos que hagan intuitiva dichas posibilidades. Por ejemplo, ante la pregunta de si es concebible un monociclo de varios kilómetros de diámetro, se puede imaginar una sucesión de monociclos, cada uno más grande que el anterior, hasta llegar al más grande de todos.

Una objeción general que se puede plantear a este tipo de experimentos mentales es que quizá no sea posible imaginar todos los rasgos relevantes para evaluar la posibilidad de ciertas situaciones, como por ejemplo la vastedad de neuronas que componen un cerebro humano. Para responder a este tipo de objeción, Chalmers propone el experimento mental de la nación isomórfica con una mente.⁴⁵ Imaginemos a los habitantes de China organizándose de tal modo que constituyan un duplicado isomórfico respecto a la organización funcional de nuestro cerebro (podrían comunicarse entre ellos con transistores, e incluso para hacer el experimento más intuitivo, podríamos pensar que los *outputs* e *inputs* vinculados a esta organización funcional podrían estar mediados y conectados con una especie de robot o humanoide, dotado de dispositivos motores). Este experimento se usa para mostrar que un artefacto como el mencionado que compartiera nuestra organización funcional carecería de experiencia consciente, ya que intuimos que carecería de fenomenología.

La premisa (1) puede entenderse también como un desafío en el que se invita a mostrar la inconcebibilidad de los zombis y que goza, a priori, de justificación. Sin embargo, ¿cuál es la noción de concebibilidad que está en juego? ¿Es la concebibilidad la ausencia de una contradicción explícita al pensar una situación, o

⁴⁴ La consideración del pansiquismo como una posibilidad inteligible se encuentra vinculada al éxito de este argumento.

⁴⁵ La primera formulación de este experimento mental se puede encontrar en Block (1978).

más bien es la imposibilidad psicológica o epistémica de reconocer la incoherencia en una situación contrafáctica? Es evidente que Chalmers no considera que los límites formales de la lógica sean las únicas restricciones a lo concebible. Como vimos al tratar el marco general de la discusión, el tipo de concebibilidad lógica se vincula a la semántica de mundos posibles, la cual es anterior, según Chalmers, a la relación lógica de deducibilidad en un sistema formal. Pero entonces, es necesario dotar de contenido la idea de una verdad necesaria en virtud de los significados de los conceptos implicados en ésta.

Por otra parte, la premisa (2) es muy sensible a la ambigüedad respecto al tipo de concebibilidad que se tiene en cuenta en (1), en la medida en que necesita un vínculo entre la concebibilidad de estados de cosas, y ciertas consecuencias modales. Parece que hay sentidos de concebibilidad que claramente no implican la posibilidad de una situación. Un ejemplo manido es la concepción de la posibilidad de que el agua no contenga hidrógeno. En principio, un hablante competente de castellano podría no saber que el agua se compone de hidrógeno, e incluso podría llegar a creer, coherentemente, que tal es el caso. Sin embargo, esto no implica que el agua podría no contener hidrógeno. Siguiendo a Kripke, una descripción más adecuada de lo que realmente concebimos coherentemente al tener en cuenta esta posibilidad, es que una sustancia semejante al agua en algunas de sus propiedades, una sustancia quizá indistinguible del agua para la mayoría de los hablantes competentes del castellano, no contiene hidrógeno. Ahora bien, cuando hablamos de la posibilidad de que determinado tipo de experiencia consciente sea instanciada con independencia de cualquier proceso físico, ¿cómo describimos adecuadamente lo que estamos concibiendo? ¿Se trata de un caso más en el que un determinado tipo de cosa coincide en sus propiedades superficiales con otra cosa, pero es distinta en su esencia?

Una buena parte del trabajo de Chalmers se encuentra dirigido, precisamente, a desarrollar un marco que dé cuenta de los nexos entre concebibilidad y modalidad. Esta es la finalidad del desarrollo de una herramienta semántica que trata de distinguir dos componentes en la determinación del significado, y que trata de dar cuenta de la especificidad semántica de los términos

que refieren a propiedades fenoménicas, así como de las paradojas aparentes que se producen entre lo a priori y a posteriori, por un lado, y lo necesario y contingente por el otro. Precisamente, por tener en cuenta dos formas de determinar el significado, dicho marco se conoce como bidimensionalismo. En la sección 4 presentaré este marco semántico con más detalle.

3.1.2. El espectro invertido

Aunque son numerosas las versiones de este argumento aparecidas en la bibliografía al respecto, el argumento básico parte de la posibilidad de que dos personas, la una respecto a la otra, sufran de una inversión de uno de los colores del espectro cromático. Imaginemos que A tiene experiencias de verde en aquellas ocasiones en las que B tiene experiencias de rojo. La pregunta es: ¿podríamos llegar a fijar un conjunto de hechos físicos que determine las experiencias cromáticas de un sujeto? Asumamos que ambos sujetos han recibido el mismo tipo de instrucción lingüística en las mismas situaciones, de manera que cualquier testimonio verbal será indistinguible en ambos casos. Por otra parte, parece que recurrir a la información que nos proporciona la neurología no nos será de gran ayuda. Una vez más resurge el problema que en el primer capítulo expusimos respecto al modelo de la co-ocurrencia: en última instancia, ¿en virtud de qué podríamos diferenciar un patrón neuronal que determine el tipo de experiencia cualitativa que el sujeto está teniendo?

Chalmers formula el argumento con un matiz distinto, de manera que lo que plantea es la concepción de un duplicado, o bien de nosotros mismos, o bien de nuestro mundo, de manera que, en cualquiera de estos casos, la experiencia fenoménica respecto al color se encuentre invertida. Ahora bien, si estas posibilidades son coherentes, entonces, el argumento continúa, los hechos respecto a la experiencia fenoménica no se encuentran determinados por los hechos físicos y, por tanto, el materialismo es falso.

De entrada, parece que la conclusión de este argumento es más débil que la que se puede extraer del argumento de los zombis. El problema de la conciencia

puede dividirse en dos cuestiones básicas: en primer lugar, ¿por qué hay experiencia fenoménica en general en un mundo físico?; en segundo lugar, ¿cómo se determina el carácter cualitativo específico de cada experiencia fenoménica dados los hechos físicos del mundo? Planteado así, el espectro invertido proporciona un argumento que apunta al escepticismo respecto a cualquier intento de determinar el carácter cualitativo de una experiencia fenoménica concreta a partir de un conjunto de hechos físicos, pero no cuestiona que, dado determinado conjunto de hechos físicos, haya, en general, experiencia consciente.

3.1.3. Asimetría epistémica

Intuitivamente, la relación con nuestros propios estados mentales es distinta respecto a la relación que tenemos con los estados mentales de los demás. Siguiendo la tradición cartesiana, se ha llegado a afirmar que cada sujeto es infalible respecto al conocimiento de sus propios estados mentales, de manera que, si a alguien le parece que está teniendo un dolor, no puede estar equivocado. Por el contrario, si de lo que se trata es de averiguar si una persona distinta de nosotros tiene un estado mental, parece que siempre es posible el error, ya que nuestro acceso a sus estados mentales siempre es mediado. Recordemos el ejemplo del médico que trata de determinar mediante un escáner si una persona tiene dolor. El problema más interesante filosóficamente no es que, en general, por cuestiones empíricas, podamos estar equivocados respecto a los estados mentales de los demás, sino que incluso, en circunstancias epistémicamente ideales, podríamos estar equivocados respecto a éstos. Como ya vimos en el capítulo anterior, esto haría de la experiencia consciente un rasgo imprevisible y sorprendente del mundo, si no fuera por el conocimiento que de esta tenemos en primera persona. Ahora bien, ¿en qué medida puede constituirse esta intuición en un argumento contra el fisicismo? Según Chalmers, a partir de nuestras intuiciones epistémicas pueden extraerse consecuencias metafísicas vinculadas a la relación de supervinencia entre los hechos físicos y mentales, de manera que:

La asimetría epistémica en el conocimiento de la conciencia deja claro que la conciencia no puede supervenir lógicamente. Si fuera lógicamente superviniente, no habría tal asimetría epistémica; una propiedad lógicamente superviniente se puede detectar fácilmente por medio de evidencia externa, y no hay un rol especial para el caso de la primera persona. (Chalmers 1996: 102)

3.1.4. El argumento del conocimiento

El otro argumento que explota la intuición de la asimetría epistémica es el conocido argumento del conocimiento de Frank Jackson (1982). Algunos de los recursos dramáticos con los que se presenta el experimento pueden variar, pero su formulación básica es la siguiente: imaginemos que Mary ha sido criada, hasta su edad adulta, en un entorno completamente carente de colores.⁴⁶ Además, Mary ha llegado a ser la mejor neurólogo del mundo especializada en la visión del color, de manera que conoce, en principio, todos los hechos neurológicos y físicos implicados en la visión de color, así como los hechos físicos que tienen que ver con los objetos vistos; sin embargo, debido a su extraña condición, Mary no sabe cómo se ve el amarillo. De hecho, podría pensarse que, si se le enseñara un plátano azul, llegaría a creer erróneamente que el color amarillo es el color azul, de modo que, si Mary llegara a ver un objeto de color, entonces aprendería un hecho nuevo del mundo. Ahora bien, si de todos los hechos físicos relevantes del mundo no se pueden deducir los hechos relativos a la experiencia consciente, en este caso a la experiencia consciente de color, entonces el materialismo es falso, ya que existiría un conjunto de hechos adicionales a los hechos físicos.⁴⁷

3.1.5. La ausencia de análisis

⁴⁶ Cómo se consigue esto es, en principio, irrelevante para el experimento. Entre los recursos dramáticos del experimento se encuentran el uso de pinturas o maquillajes, de monitores en blanco y negro, etc.

⁴⁷ Chalmers vincula este experimento al argumento de Thomas Nagel en *What is it Like to Be a Bat?*, a pesar de que el mismo Jackson, en el artículo original, considera que ambos argumentos deben distinguirse.

Finalmente, el último argumento propuesto por Chalmers se vincula de una manera directa al análisis de nuestros conceptos. Si recordamos el marco en el que Chalmers sitúa el estudio de la mente, el análisis funcional de los estados mentales es exitoso, en principio, para comprender todos aquellos aspectos de la mente vinculados al uso y recepción de la información con la finalidad de regir y controlar el comportamiento. Sin embargo, en el caso de la experiencia consciente, este análisis no sólo no existe, sino que no tiene sentido tratar de llevarlo a cabo: “Aunque los estados conscientes pueden desempeñar varios roles causales, no se definen por sus roles causales. Más bien, lo que los hace conscientes es que tienen un sentir fenoménico, y este sentir no es algo que se pueda definir funcionalmente” (Chalmers 1996: 105).

Ahora bien, ¿es esto un argumento contra la tesis ontológica del materialismo? ¿No se está hablando, exclusivamente, en términos de explicación reductiva? La cuestión es que, precisamente, el hecho de que no exista un análisis funcional de la experiencia consciente tiene consecuencias directas para la tesis materialista. La articulación que Chalmers lleva a cabo del materialismo se vertebra en torno a la noción de supervinencia; sin embargo, en sí misma, la relación de supervinencia no es sino una relación neutral respecto a diferentes concepciones metafísicas. Lo que hace realmente importante a la relación de supervinencia es su vínculo con la tesis de que las propiedades mentales son propiedades funcionales realizadas por propiedades físicas, de manera que, si una propiedad mental no puede analizarse, por principio, en dichos términos, se pierde la justificación de la relación de supervinencia y, por ende, de la tesis materialista.

3.2. Consideraciones generales sobre los argumentos

Los argumentos presentados siguen, según Chalmers, tres estrategias diferentes. El argumento de los zombis y el del espectro invertido apelan a lo que es concebible en situaciones contrafácticas en las que los hechos físicos no varían. Tanto el argumento de la asimetría epistémica como el argumento del conocimiento se basan en que la forma de conocer los hechos físicos es diferente del conocimiento

que tenemos de la experiencia consciente. Finalmente, el argumento de la ausencia de análisis se sirve del mismo concepto de experiencia fenoménica para concluir su irreductibilidad, aunque quizá considerar que hay diferentes estrategias oscurece el hecho de que todos los argumentos comparten algunas intuiciones básicas. Por ejemplo, es difícil sustraerse a la idea de que los argumentos que apelan a la concebibilidad se nutren de la intuición del acceso asimétrico a nuestros estados mentales y, del mismo modo, la conclusión de que hay una falta de análisis en términos funcionalistas para la fenomenología no es independiente de dicho acceso.

El argumento de los zombis expresa de un modo más claro por qué la experiencia fenoménica constituye un problema para el fisicismo y muestra con claridad que tiene que haber un nexo entre lo concebible y la metafísica. Veamos de nuevo la estructura del argumento:

- (1) Los zombis son concebibles.
- (2) Si los zombis son concebibles, entonces son lógicamente posibles.
- (3) Si los zombis son lógicamente posibles, el materialismo es falso.
- (4) Por lo tanto, el materialismo es falso.

Una objeción fácil a este argumento es considerar que cualquier intento de articular una vía entre concebibilidad y posibilidad es sospechoso, y podría resultar que, simplemente, hay verdades metafísicamente necesarias, pertenecientes a un subconjunto de verdades necesarias independiente de la necesidad lógica. Pero si la única razón para postular dichas identidades es el problema que plantea la experiencia fenoménica, resulta una maniobra *ad hoc*. Independientemente, habría que dar una explicación plausible de por qué la conciencia es especial. Además, este carácter *ad hoc* de dicha respuesta tiene que enfrentarse, además, al problema de la concepción general de la modalidad y su inteligibilidad. Chalmers lo expresa del siguiente modo:

La respuesta corta a esta objeción es que no hay razón para creer que esta modalidad exista. Tales “necesidades metafísicas” pondrán restricciones al espacio de mundos físicos que son brutos e inexplicables. Puede ser razonable permitir la existencia de hechos brutos e inexplicables sobre *nuestro* mundo, pero la existencia de dichos hechos sobre el espacio de mundos posibles sería bastante extraña. El reino de lo posible (como opuesto al reino de lo natural) no tiene espacio para este tipo de restricción arbitraria. (Chalmers 1996: 137)

Chalmers no explica por qué este tipo restricción resulta extraño; al fin y al cabo, después de la obra de Kripke no es razonable asumir, por defecto, qué es o no posible. En caso de que éste fuera un dilema real entre rechazar la verdad del fisicismo o comprometerse con la existencia de necesidades brutas a posteriori, todavía no se han dado razones independientes para decantarse por uno u otro cuerno del dilema. Si los argumentos a favor del fisicismo funcionan, no está claro por qué las consideraciones a favor de la racionalidad de nuestras intuiciones modales tendrían que prevalecer sobre unos argumentos sólidos a favor del fisicismo (véase Levine 2001).

La posibilidad de los zombis resulta, de entrada, incluso extravagante, de manera que, por ejemplo, Jaegwon Kim (2005) rechaza la fantasía de los zombis como un argumento convincente que muestre la irreductibilidad de la experiencia consciente y apunta a la concebibilidad del espectro invertido como un argumento más intuitivo. No obstante, no está claro en qué medida el argumento del espectro invertido es realmente diferente respecto al argumento de los zombis. Volviendo al argumento de la concebibilidad, el condicional expresado por (3) podría sustituirse por el siguiente condicional:

(3') Si el espectro invertido es lógicamente posible, el materialismo es falso

Y mediante (3') podría establecerse igualmente la falsedad del materialismo, en el sentido definido por Chalmers. Sin embargo, ¿es coherente aceptar el argumento del espectro invertido y rechazar, a la vez, el argumento de los zombis?

Si bien para Chalmers esta no es una cuestión demasiado importante, puesto que su objetivo es mostrar que la experiencia consciente constituye un contraejemplo al fisicismo, es interesante analizar las relaciones entre el antecedente y el consecuente para comprender mejor las intuiciones en las que se basa el argumento de la concebibilidad. Veamos cuál es la razón que, según Chalmers, apoya la posibilidad lógica del espectro invertido: “Nada en la neurofisiología dictamina que un tipo de procesamiento debería estar acompañado por experiencias rojas en lugar de amarillas” (Chalmers 1996: 100).

En primer lugar, no está claro si esta es una razón que explica la posibilidad del espectro invertido, o simplemente es una constatación de la misma intuición. Al fin y al cabo, decir que nada en la fisiología determina un veredicto respecto a la experiencia fenoménica, si no se explica mejor, no es muy distinto de concebir una situación en la que, dado cierto conjunto de propiedades neurofisiológicas, no se dan las correspondientes propiedades fenoménicas. Además, si éste es el principio que se encuentra detrás de la plausibilidad del espectro invertido, tampoco está muy claro por qué debería tener una formulación tan restringida. Si no hay razón para que un determinado estado neurofisiológico vaya acompañado de una experiencia de amarillo en lugar de una de rojo, ¿por qué tendría que estar acompañada de experiencia alguna?

Parece que la plausibilidad del experimento depende más de cierta retórica vinculada al ejemplo del color que no funciona tan bien si variamos el ejemplo. ¿Por qué no se habla del experimento mental de una inversión sinestésica? ¿O de un sujeto tal que cuando los demás experimentan una sensación de rojo, él experimenta un sonido de trompeta y, aun siendo esto así, su comportamiento y sus reacciones son indistinguibles de las de un individuo normal? Incluso la inversión exclusiva de un solo color del espectro cromático resulta problemática, ya que, como resultado de dicha versión, las relaciones entre los distintos colores se alterarían y no resultaría plausible la idea de que una diferencia en la experiencia fenoménica de color no sea conductualmente detectable; en todo caso, tendría que haber una inversión completa del espectro cromático, de manera que las relaciones entre los distintos tonos del espectro cromático se mantuviera constante (véase Shoemaker

1982). Pero el experimento resulta todavía menos plausible cuando se imagina una inversión entre dos sensaciones del mismo tipo que sean muy diferentes cualitativamente. Imaginemos, por ejemplo, una persona que goza de sensaciones gustativas ligeramente dulces cuando los demás tienen sensaciones muy intensas de algo salado y viceversa, que además mantiene un comportamiento indistinguible del resto de las personas; ¿diríamos en tal caso que este tipo de inversión es posible?

Si se sugiere que el caso del color es diferente, entonces hay que enfrentarse a varias dificultades importantes. En primer lugar, en caso de que sea un defensor de la irreductibilidad de la experiencia consciente, habrá que reconocer que hay ciertas restricciones al tipo de experiencias conscientes que se pueden tener dadas ciertas circunstancias, de manera que el problema se limitaría notablemente. En segundo lugar, habría que explicar por qué el caso del color es diferente. En este sentido, hasta Chalmers reconoce que hay cierta evidencia empírica que sugiere la imposibilidad de la inversión del espectro cromático; por ejemplo, algunos autores (véase Harrison 1973 y Hardin 1987) señalan que hay ciertas asimetrías en los colores que hacen imposible la inversión: a ciertos colores se les asocia cierta calidez o frialdad, y no todos los colores tienen el mismo número de tonos distinguibles.

Como respuesta, Chalmers señala que no hay nada incoherente en imaginar una “fenomenología fría” acompañada de reacciones cálidas. Por otra parte, aun si la asimetría de las series tonales de distintos colores no hace posible la inversión perfecta de un color, podemos pensar en criaturas cuyos espacios cromáticos entre colores son perfectamente simétricos. Esta manera de responder es muy indicativa, no sólo de cierta concepción de la experiencia consciente, que en parte hemos analizado en el Capítulo 2, sino también del funcionamiento de la dialéctica que conduce al problema de la conciencia. Parece que, ante toda propuesta de anclar la conciencia al mundo, es posible factorizar la experiencia consciente de manera que siempre resta un componente para el cual no hay un vínculo necesario con el resto de las cosas. En este caso, la retórica del argumento vinculado al espectro invertido puede comprenderse como ligada a ciertas características especiales de

la experiencia cromática. Si comparamos una experiencia de rojo con una experiencia de dolor, parece que hay un contraste notable en la medida en que el dolor parece fuertemente vinculado a otros estados mentales y a reacciones corporales; en gran número de situaciones, un determinado gesto o movimiento es un signo bastante fiable de la presencia de un dolor, así como de su cualidad. Sin embargo, ¿qué tipo de expresión corporal puede ayudarnos a determinar si una persona está teniendo una experiencia cromática?

Es interesante observar que este modo de plantear la pregunta contiene ya una manera específica de comprender tanto la relación entre lo mental y lo corporal, como una metafísica de lo mental. El modelo de la co-ocurrencia vuelve a aparecer, en la medida en que se presupone que un estado mental posee unas condiciones de individuación independientes respecto a cualquier manifestación externa, de manera que entre ambos hay, a lo sumo, una relación empírica. Quizá, tomar como básico el experimento del espectro invertido dota al problema de la conciencia de cierta retórica adicional de la que carece la concebibilidad de los zombis; sin embargo, no es evidente que ambos experimentos no sean equivalentes en sus consecuencias lógicas.

Los argumentos que parten de premisas epistémicas tratan, igualmente, de extraer una conclusión metafísica, apelando al modo especial que tenemos de conocer nuestros propios estados mentales. No voy a decir mucho más del argumento de la asimetría epistémica. Si el argumento que presenté al final del Capítulo 2 funciona, no es evidente que el reconocimiento de una asimetría epistémica respecto a lo mental permita concluir que no hay vínculos conceptuales entre la experiencia consciente y criterios públicos de identidad de los estados mentales conscientes. En todo caso, la asimetría epistémica constituye un problema filosófico interesante por sí mismo, pero no es evidente que nos permita extraer conclusiones metafísicas. En el Capítulo 6 hablaré sobre la relación entre la asimetría epistémica y el problema de la conciencia.

El argumento del conocimiento se sirve también de una situación hipotética, si bien trata de ilustrar la diferencia entre cómo conocemos los hechos relativos a nuestra propia experiencia consciente y los hechos físicos. Sin embargo, no está

muy claro si el argumento se sirve de una noción de hecho adicional relevante que permita establecer una conclusión contra el fisicismo. Recordemos que el fisicismo, y especialmente como lo entiende Chalmers, no se compromete con la idea de que todos los hechos sean físicos, sino que, más bien, los hechos físicos del mundo fijan el resto de hechos de un modo lógicamente necesario. En general, los argumentos que parten de premisas epistemológicas y tratan de llegar a conclusiones metafísicas suelen encontrarse sujetos a un tipo de objeción general vinculada a la idea de que el conocimiento incluye, ineludiblemente, un elemento perspectivista. En este sentido, alguien puede saber que en Japón se ha producido un terremoto porque se encontraba allí en ese momento, o bien porque lo ha visto en un informativo; por supuesto, hay grandes diferencias entre las dos formas de llegar a saber sobre el terremoto, pero hay un único hecho, a saber, que se ha producido un terremoto en Japón. Un mismo hecho se puede presentar de distintas maneras.⁴⁸ Del mismo modo, quizá en el caso de Mary todo lo que haya sea el conocimiento del hecho respecto a cómo se siente una experiencia de rojo, pero conocido desde distintas perspectivas, de modo que hay un único hecho físico. El problema de esta objeción es que el reconocimiento de la existencia de distintos modos de presentación hace que el argumento del conocimiento se reproduzca respecto a un hecho adicional que vincula los medios de presentación, por ejemplo: esta experiencia que tengo en este momento es idéntica al tipo de actividad neuronal X producida por esa longitud de onda L. Pero, ¿de qué manera podría haber conocido Mary este nuevo hecho si no es teniendo la experiencia cromática apropiada?⁴⁹ Lewis (1990) se niega que Mary adquiera el conocimiento de un hecho, sino que más bien posee una nueva habilidad; es capaz, por ejemplo, de discriminar sombras de color cuando le son presentadas. Por otra parte, Dennett (1991) ha señalado que el tipo de experimento mental propuesto por Jackson no permite extraer fiablemente el tipo de conclusión que pretende, en la medida en que exige imaginar unas

⁴⁸ Para una formulación de este argumento véase Moya (2004).

⁴⁹ Algunos filósofos tratan de dar una respuesta al problema del hueco explicativo situando la cuestión en el terreno de nuestros propios conceptos, y no en la ontología. En este sentido, se comprometerán con la existencia de unos conceptos específicos vinculados a la experiencia fenoménica y una de sus características especiales será que carecen de un modo de presentación.

condiciones muy difíciles de abarcar, y no se puede concebir de una manera clara lo que implica conocer la totalidad de hechos físicos relevantes implicados en la visión del color. Dadas dichas condiciones poco fiables, un resultado del experimento podría ser que Mary reconociera sin problemas y sin sorpresa alguna el tipo de color que se le presenta, y que, al presentársele un plátano azul, fuera capaz de identificar el intento de engaño. No obstante, Chalmers podría argumentar que el peso de la prueba está en el otro lado, de manera que, si es imposible la concepción coherente de una determinada posibilidad, esto debería poder mostrarse de algún modo.

Por el momento, no daré una respuesta concluyente a este tipo de argumento, pero cabe señalar que resulta un rasgo significativo del experimento que una vez más se recurra a la experiencia cromática para extraer conclusiones que afectan a la experiencia consciente en su totalidad. Incluso podría haberse barajado un caso real en el que una persona tiene una sensación por primera vez, aunque previamente tenía ya un conocimiento de los hechos vinculados a dicha sensación. Por ejemplo, podríamos pensar en un médico especialista en un determinado tipo de enfermedad que, por otra parte, no la ha experimentado en primera persona. Dicha enfermedad se puede manifestar como una serie de dolores muy específicos y muy concretos (podemos imaginar que es una serie de hormigueos en los brazos, acompañada de un dolor punzante que se difumina desde el estómago hasta la espalda, etc.) y podemos pensar que la enfermedad se manifiesta en otros síntomas bastante concretos, quizá como un cambio en el color de la piel. Supongamos que un lego en medicina sufre esa enfermedad y, consiguientemente, el tipo de dolor específico descrito: ¿diríamos que conoce un hecho adicional respecto a la enfermedad que no conozca el médico especialista? ¿Y si el médico contrajera la enfermedad adquiriría el conocimiento del hecho adicional de la misma manera, sin que el conocimiento previo le permita identificar con más facilidad y precisión el tipo de sensaciones que tiene? En este caso, las intuiciones no son tan claras. Por supuesto, se podría responder fácilmente que estamos hablando de otra cosa, ya que no se trata del caso equivalente de Mary traspuesto a la sensación del dolor; de hecho, el médico ya está familiarizado con

sensaciones dolorosas, a pesar de que no haya tenido los dolores específicos de la enfermedad en cuestión, y su situación sería semejante, más bien, a la de una persona que no conoce determinada tonalidad de un color. Sin embargo, si intentamos realizar el mismo experimento con el dolor nos encontramos con grandes problemas para concebirlo. De hecho, podemos ahorrarnos muchos recursos dramáticos, ya que existe un trastorno que consiste en la incapacidad para tener sensaciones dolorosas. Imaginemos ahora un científico semejante en sus conocimientos científicos a Mary, pero especialista en el dolor que, además, padece de ese extraño trastorno que no le permite sentir dolor. La cuestión es la siguiente: para que el experimento llegue a ser exitoso tenemos que ser capaces de concebir coherentemente que el científico conoce todos los hechos físicos y neurofisiológicos implicados en el dolor; sin embargo, ¿puede ser este el caso de una persona que no siente el dolor? Por ejemplo, ¿puede decirse que una persona que no siente dolor conoce todos los hechos relevantes de la física respecto a la solidez de los cuerpos o al calor? ¿Conoce el mundo en la misma medida que los demás? Tener una concepción del mundo es en parte tener una concepción de aquello que no es uno mismo; ahora bien, ¿esta idea de lo que no es uno mismo permanece inalterada a pesar de que uno pueda ser capaz de mutilar su propio cuerpo de la misma manera que desmonta un artefacto? Por supuesto, podrían conocerse todos los hechos del mundo en el sentido de memorizar todos los enunciados de los manuales de física o de neurofisiología, pero alguien podría memorizar dichos enunciados, e incluso, sobre el papel, mostrar que tiene cierta comprensión de lo que ha memorizado y, sin embargo, desconocer que dichos enunciados hablan del mundo real o de su propio cerebro. Esta sería una manera muy rara de conocer los hechos del mundo y también de concebir la relación de la ciencia con el mundo. Implícitamente, el argumento del conocimiento asume cierta concepción de lo que significa conocer un hecho y genera la intuición de que hay algo que Mary no sabe y que es central en el valor de la vida humana. Una pregunta interesante es por qué a Mary, con todo su conocimiento científico, se le escapa algo que resulta central a la vida humana. En el Capítulo 6 hablaré de la ciencia de la conciencia y de los

límites que dicha ciencia tiene en el marco en el que se plantean las discusiones actuales.

Finalmente, el argumento de la ausencia de análisis parece merecer una categoría aparte. Este argumento se encuentra ligado a la concepción funcionalista de la mente, de manera que una explicación reductiva de la experiencia consciente solo sería posible si existiera un análisis funcionalista de la experiencia fenoménica, pero este análisis no existe. Este argumento introduce algunos supuestos controvertidos. Por ejemplo, presupone que, si la experiencia consciente pudiera someterse a algún tipo de análisis, éste tendría que ser, exclusivamente, funcionalista. Sin embargo, hablar de experiencia consciente y, a la vez, proponer un proyecto que trate de reducirla al desempeño de un rol causal, no parece que sea una empresa que pueda iniciarse, para después descubrir, empíricamente, que es irrealizable. Tal y como lo plantea Chalmers, hablar de dicho proyecto es una contradicción, ya que comprender el *explanandum* equivale, en parte, a comprender que cualquier explicación en términos de roles causales está excluida. Por otra parte, la validez de este argumento se encuentra condicionada a que el éxito de la explicación de la experiencia consciente dependa del éxito de la concepción funcionalista de la mente (para una posición de este tipo véase Kim 2005), algo lejos de estar firmemente establecido, incluso por razones independientes de la discusión sobre la conciencia.

De las consideraciones de esta sección, se pueden decir algunas cosas generales respecto a los argumentos. En primer lugar, parece que los argumentos no son conceptualmente independientes, aunque su fuerza retórica varíe en función del ejemplo al que recurren en cada caso, de modo que, con ejemplos de diferentes sensaciones, los argumentos no resultan igualmente intuitivos ni convincentes. En segundo lugar, los argumentos, tal y como se presentan, tienden a situar la carga de la prueba en el otro lado. Algunos de los argumentos recurren a la concebibilidad de ciertas situaciones que no son evidentemente concebibles; sin embargo, mientras no se muestre que dichas posibilidades no están excluidas lógicamente, parece que los argumentos son atractivos, o al menos generan una perplejidad que pone en jaque el marco fisicista. En el caso del argumento de la asimetría

epistémica, el hecho de que exista la asimetría ya supone una dificultad para el fisicismo. Pero si estos argumentos no cuentan con un movimiento dialéctico adicional que permita vincular lo concebible con lo que es posible lógica y metafísicamente, parece que, a lo sumo, logran establecer una conclusión escéptica o la existencia de una dificultad especial para el fisicismo. No obstante, como ya vimos a propósito del argumento causal a favor del fisicismo, puede haber razones independientes para defender el materialismo, y quizá todo lo necesario para defenderlo plausiblemente sea mostrar que las propiedades fenoménicas de la experiencia consciente pueden ser, en última instancia, propiedades físicas. En este sentido, es posible que el tipo de problemas que se plantean respecto a la conciencia no sea tan especial, sino que más bien el hueco teórico con el que se encuentra Chalmers se genere a propósito de otras identidades respecto a las que no parece haber duda razonable, como en el caso de la identidad entre el agua y el H₂O. Después de la obra de Kripke, y el consecuente divorcio entre lo *a priori* y lo necesario, cualquier argumento que trate de extraer conclusiones a partir de las relaciones entre ambos necesitará de un marco adicional que permita dar cuenta de las distintas intuiciones modales en conflicto. Si una identidad paradigmática, desde el punto de vista científico, como lo es la del agua y H₂O es un ejemplo de una identidad necesaria que, además, se ha establecido a posteriori, ¿qué nos impide concluir que, del mismo modo, la experiencia fenoménica es un determinado tipo de estado cerebral?⁵⁰

Hasta ahora, solo he explicado de qué manera articula Chalmers su concepción del materialismo en torno a la noción de supervinencia lógica, en el sentido vinculado a la semántica de mundos posibles, pero todavía no se ha dicho nada respecto al modo general que tenemos de evaluar un determinado estado de cosas en un mundo posible, o a la justificación que puedan tener nuestras intuiciones modales en dicho marco. Esta consideración es importante en la medida en que Chalmers necesita establecer un vínculo entre concebibilidad y metafísica que le permita extraer una conclusión antimaterialista. Además, una objeción que

⁵⁰ Si bien la distinción de Kripke puede servir para justificar un argumento a favor del materialismo, el mismo Kripke desarrolló un argumento antimaterialista en relación a determinados estados mentales, como veremos en este mismo capítulo.

se plantea de un modo inmediato a los argumentos planteados es que no importa lo que concibamos o creamos concebir respecto a la conciencia, ya que empíricamente podríamos llegar a saber que la experiencia consciente es idéntica a un conjunto de propiedades físicas, y en este caso sería así de un modo necesario; es necesario un movimiento dialéctico adicional, por parte de Chalmers, que le permita bloquear este tipo de objeción. Como ya adelanté, Chalmers desarrollará un modelo semántico con el fin de dar cuenta de las ambigüedades aparentes que aparecen al aplicar conceptos en el espacio de mundos posibles. A continuación, describiré los rasgos generales de este marco semántico y argumentaré que difícilmente puede ayudar a comprender mejor la experiencia consciente.

3.3. Concebibilidad, posibilidad lógica y necesidad⁵¹

Puesto que para que el argumento antimaterialista de Chalmers funcione es necesario que haya un vínculo entre concebibilidad y posibilidad, una formulación más completa del argumento debería hacer explícito dicho vínculo (en lo que sigue, me serviré de Chalmers 2002b y Chalmers 2010). Asumamos que F es el conjunto de los hechos físicos y macrofísicos del mundo real y C el conjunto de los hechos relativos a la experiencia consciente en el sentido fenoménico, también en el mundo real. Una versión inicial del argumento, posiblemente la más sencilla, podría ser la siguiente:

- (1) $F \ \& \ \sim C$ es concebible.
- (2) Si $F \ \& \ \sim C$ es concebible, $F \ \& \ \sim C$ es metafísicamente posible.
- (3) Si $F \ \& \ \sim C$ es metafísicamente posible, el materialismo es falso.

⁵¹ En este apartado, presento las líneas generales del marco bidimensionalista desarrollado por Chalmers. Las intuiciones básicas de las que parte se encuentran en Kripke (1980). Se puede encontrar una exposición general de la semántica bidimensionalista y sus aplicaciones generales en Schroeter (2012). El marco bidimensionalista es una herramienta formal que puede tener diversas aplicaciones. Una de las aplicaciones más aceptadas es la que da cuenta del funcionamiento de expresiones deícticas, véase Kaplan (1989).

(4) Por lo tanto, el materialismo es falso.

Si comparamos este argumento con la versión formal del argumento de los zombis que presenté en la sección anterior, podemos observar algunas diferencias. En primer lugar, en esta formulación se especifica que la posibilidad a concebir es un mundo en el que se mantienen los hechos físicos reales sin que vayan acompañados de los mismos hechos relativos a la experiencia consciente. Esta formulación es más amplia que la de los zombis y sintetiza la intuición general que hay tanto en los argumentos de los zombis y del espectro invertido como, posiblemente, en el argumento del conocimiento. Por otra parte, ya no se habla de posibilidad lógica, sino de posibilidad metafísica, dado que la posibilidad lógica a la que se refiere Chalmers hace referencia a la semántica de mundos posibles y no a la deducibilidad en un sistema de lógica formal. Al hablar de este tipo de posibilidad lógica, Chalmers responde directamente a las objeciones que se le pueden plantear desde un tipo de posibilidad que no es determinable a priori. Como veremos a continuación, según Chalmers, no hay una posibilidad metafísica que sea completamente independiente de cómo evaluamos nuestros conceptos a priori y, para mostrarlo, desarrollará un marco semántico que le permitirá establecer vínculos estrictos entre lo concebible y lo metafísicamente posible, dando una justificación para (2).

La idea de concebibilidad se encuentra cargada de un matiz psicológico que la hace, en principio, inadecuada para extraer conclusiones metafísicas, ya que resulta difícil establecer un acuerdo universal respecto a lo concebible, y si no hay un acuerdo respecto a las intuiciones de partida, difícilmente podría avanzar una discusión iniciada en dichos términos. Chalmers es consciente de esta dificultad, y por ello traza una distinción entre concebibilidad *prima facie* y concebibilidad ideal. La concebibilidad *prima facie* se encuentra vinculada a un sujeto particular, dada cierta consideración inicial, por parte del sujeto, de la noción en cuestión; por el contrario, la concebibilidad ideal hace referencia a un sujeto ideal, libre de las limitaciones cognitivas empíricas de un sujeto real. Lo que presupone esta

concepción es que la falta de convergencia en nuestras intuiciones modales se debe, exclusivamente, a nuestras limitaciones cognitivas empíricas. Un ejemplo que hace plausible esta distinción es nuestra limitación para el cálculo matemático: podemos concebir como falso cierto teorema no evidente, cuando en realidad no lo es. Sin embargo, un ser idealmente racional podría tener acceso cognitivo a una demostración del mismo, o incluso a todos los casos posibles en los que dicho teorema se aplica.

Otra distinción previa de Chalmers incluye dos formas de entender la concebibilidad, de un modo negativo o de un modo positivo. Un determinado estado de cosas S es concebible negativamente si no se excluye a priori, y es positivamente concebible si puede imaginarse coherentemente una situación en la que S es el caso. Pensemos, por ejemplo, en la conjetura de Goldbach: todo número par mayor que 2 se puede escribir como la suma de dos primos. Aunque esta proposición ha sido verificada hasta el momento, todavía no existe una demostración de la misma. Ahora bien, ¿significa esto que la falsedad de la conjetura de Goldbach es concebible? Y si así es, ¿es posible la falsedad de la conjetura? Parece que de un modo negativo es concebible, ya que no se puede excluir, en el momento actual, a priori. No obstante, es importante señalar que las dos distinciones presentadas pueden combinarse entre sí dando lugar a distintos tipos de concebibilidad. En este sentido, la falsedad de la conjetura de Goldbach es *prima facie* negativamente concebible, pero no está claro que sea idealmente concebible (especialmente negativamente, aunque tampoco positivamente), y no parece que sea, en ningún caso, positivamente concebible.

Al margen de estas distinciones, el problema central al que se tienen que enfrentar los argumentos presentados es que la concebibilidad parece un punto de partida demasiado endeble para extraer conclusiones metafísicas. Desde que Kripke planteara la posibilidad de que hay verdades necesarias conocidas *a posteriori* (véase Kripke 1980), un argumento que trate de determinar lo que es metafísicamente posible a partir de escenarios diseñados y evaluados al margen de consideraciones empíricas parece inadecuado para derivar conclusiones metafísicas, y aun si tenemos en cuenta la versión ideal de la concebibilidad, todavía

tendría que argumentarse a favor de la racionalidad de las nociones modales, excluyendo las necesidades brutas, ya que, de otro modo, los argumentos presentados no resultarían fiables como puente al espacio de la posibilidad metafísica. Chalmers se sirve de un modelo semántico que tiene en cuenta dos formas de evaluar nuestras intuiciones modales a través del espacio de mundos posibles, de ahí que dicho marco se conozca como bidimensionalista. Veamos en qué consiste.

Para empezar, tomemos un ejemplo paradigmático de lo que se considera una verdad necesaria *a posteriori*:

(a) El agua es H₂O

Si tenemos en cuenta las distinciones introducidas, ¿podemos concebir en algún sentido que (a) sea falsa? Parece que *prima facie* sí, ya que no todo sujeto que use correctamente el término “agua” tendría por qué estar familiarizado con la composición molecular del agua. Por otra parte, ¿qué ocurre en este caso si tenemos en cuenta el sentido ideal de concebibilidad? ¿No es evidente que no podemos llegar a saber (a) *a priori*? Pensemos que alguien imagina que el agua no es H₂O, ¿está concibiendo realmente que el agua no es H₂O? Un ejemplo clásico de la bibliografía sobre la teoría de la referencia es el célebre experimento mental de la Tierra Gemela (véase Putnam 1975), donde un planeta idéntico a la Tierra en casi todos los aspectos se diferencia de esta, exclusivamente, en que el líquido inodoro, incoloro e insípido que llena mares y océanos tiene una composición química diferente: en lugar de ser H₂O es XYZ. En este experimento mental, ¿es la Tierra Gemela un estado de cosas en el que el agua no es H₂O? La explicación que Kripke da de estos casos es que realmente no se está concibiendo la falsedad de (a), sino más bien que un líquido semejante en sus propiedades superficiales al agua no es H₂O. Ahora bien, ¿supone esta respuesta un problema para la concepción de Chalmers? La interpretación general de Chalmers es que cuando evaluamos la aplicación de un concepto a través del espacio de los mundos posibles se puede hacer de dos maneras distintas. En primer lugar, los mundos posibles

pueden ser considerados como reales, y se trata de mundos en los que se ha indicado un centro desde el que se evalúa la aplicación del término en cuestión; para definir dicho centro se señalará un individuo y un tiempo determinado. Según esta forma de evaluar la aplicación posible de un término, la pregunta que se lleva a cabo en la Tierra Gemela respecto a si el agua es H₂O es negativa, ya que:

(b) El líquido incoloro, inodoro e insípido que cubre mares y océanos y cae del cielo es XYZ.

La necesidad de este análisis deriva de la idea de que, para poder llegar a descubrir la composición molecular del agua, hay que tener una concepción previa de la sustancia cuya composición se va a determinar. Esto es lo que Chalmers denomina una intensión primaria y su papel consiste en determinar de qué manera la referencia depende de cómo es el mundo externo y, por ende, ella misma es previa lógicamente a cómo es el mundo externo de hecho, siendo, por tanto, *a priori*. Ahora bien, dado que el mundo es de una determinada manera y que en la Tierra el líquido incoloro, inodoro, ... etc. es H₂O, quedando de este modo la referencia del término "agua" fijada, podemos considerar otros mundos posibles como contrafácticos, y señalar a la misma sustancia en todos los mundos posibles. Esto es lo que Chalmers denomina intensión secundaria, y se determina *a posteriori*. Su nexa con la intensión primaria es fundamental, en la medida en que la intensión secundaria se determina, previamente, evaluando la intensión primaria en el mundo real. Cuando Kripke afirma que (a) es una verdad necesaria, está teniendo en cuenta la intensión secundaria, pero al contemplar la posibilidad de un mundo en el que el líquido superficialmente semejante al agua no es H₂O, está reconociendo, de acuerdo con la interpretación de Chalmers, cierto nexa entre concebibilidad y posibilidad que se vincula a las intensiones primarias.

Estos dos tipos de intensión se corresponden, a su vez, en dos formas adicionales de concebibilidad, de modo que S es primariamente concebible si es concebible que S sea realmente el caso (intensión primaria), y S es secundariamente concebible si es concebible que S podría haber sido el caso

(intensión secundaria).⁵² En última instancia, la intuición que trata de atrapar el modelo bidimensionalista es que un usuario competente de una lengua es capaz de evaluar la aplicación de conceptos a partir de unas circunstancias determinadas dadas, sin necesidad de recurrir a información empírica adicional, de modo que no es necesario que dicho usuario asocie una determinada descripción a un concepto para poder aplicarlo. Chalmers identifica los dos tipos de intensiones con funciones. En el caso de la intención primaria, se trata de una función que hace corresponder mundos posibles considerados como reales con extensiones, y en el caso de la intención secundaria se trata de una función que hace corresponder mundos posibles considerados como contrafácticos, ya que previamente se ha fijado qué referente designa la intención primaria dadas las circunstancias del mundo real, con extensiones. Cuando de lo que se trata es de evaluar las intensiones primarias y secundarias de enunciados, se obtienen dos funciones que hacen corresponder mundos posibles con valores de verdad.⁵³

Este modelo da cuenta de enunciados como (a) del siguiente modo: según la intención primaria, el agua es H₂O en la Tierra, y XYZ en la Tierra Gemela, de manera que (a) sería verdadero en la Tierra, pero falso en la Tierra Gemela; ahora bien, dado que en el mundo real el agua es H₂O, hay un sentido según el cual el agua es H₂O en todo mundo posible, y esta intuición viene recogida por la intención secundaria, de manera que, según esta, (a) es verdadero en todo mundo posible.

Veamos ahora cómo se aplica esta distinción al problema de la conciencia. Tomemos una hipótesis materialista tentativa para un tipo de propiedad fenomenológica como el dolor:

(c) El dolor es actividad de las fibras nerviosas C.

⁵² A su vez, esta distinción adicional puede ser combinada con las otras, dando lugar a distintas formas de concebibilidad primaria y secundaria.

⁵³ La función que constituye la intención primaria puede incluir un argumento adicional que determine el punto del mundo desde el que se evalúa la intención, con el fin de evitar casos de ambigüedad cuando hay más de un candidato que puede considerarse la extensión de la función.

Un antagonista de Chalmers podría defender la necesidad de un enunciado como (c) afirmando que ésta es una identidad descubierta empíricamente, que de ningún modo puede cuestionarse desde consideraciones puramente apriorísticas. No obstante, si aceptamos el marco bidimensionalista, ¿puede hacerse un análisis de (c) equivalente al de (a)? La respuesta de Chalmers es que no. Lo que los argumentos de la sección anterior muestran, en caso de funcionar, es que, de acuerdo con la intensión primaria, hay mundos posibles físicamente idénticos al nuestro que carecen de experiencia consciente. Si el caso de (a) fuera equivalente al de (c), entonces, según la intensión primaria de agua, sería posible concebir un mundo físicamente idéntico al nuestro pero carente de agua y esto equivaldría a decir que hay mundos posibles físicamente idénticos al nuestro que carecen de una sustancia que llene los océanos y lagos, y caiga del cielo, lo cual resulta completamente contraintuitivo. Parece claro, entonces, que hay una asimetría clara entre “dolor” y términos de género natural como “agua” y “calor”.

Este argumento asume vínculos entre lo concebible y lo posible, pero ¿cuáles son estos vínculos? Volviendo al análisis de Kripke, parece que algo primariamente concebible no nos permite concluir que sea metafísicamente posible. Concebir primariamente que el agua no sea H_2O , no implica que el agua pueda no ser H_2O . Sin embargo, puede decirse que cuando concebimos la Tierra Gemela como real, concebimos que el líquido acuoso que llena mares y océanos no es H_2O , pero no podemos concluir que existe la posibilidad metafísica de que el agua no sea H_2O , dado que en el mundo real el agua es H_2O . Sin embargo, podemos decir que la concebibilidad primaria nos da acceso a un mundo posible tal que, si fuera real y el líquido que llena mares y océanos en dicho mundo fuera XYZ, entonces el agua no sería H_2O . A este tipo de posibilidad la llamaremos primaria, de manera que es primariamente posible que el agua no sea H_2O . Cuando. Por otra parte, si de lo que se trata es de evaluar la intensión secundaria de un enunciado y concluimos que el enunciado es verdadero en algún mundo posible, entonces diremos que dicho enunciado es secundariamente posible. Siguiendo con el mismo ejemplo, no es secundariamente posible que el agua no sea H_2O .

Este análisis nos puede llevar a concluir que no hay una relación entre lo primariamente concebible y lo secundariamente posible. Dadas las relaciones entre concebibilidad y posibilidad que hemos descrito, en el mejor de los casos, este es el argumento que podríamos construir:

- (1) $F \ \& \ \sim C$ es concebible.
- (2) Si $F \ \& \ \sim C$ es concebible, $F \ \& \ \sim C$ es primariamente posible.
- (3) Si $F \ \& \ \sim C$ es primariamente posible, el materialismo es falso.
- (4) Por lo tanto, el materialismo es falso.

El problema con este argumento es que, si bien (1) y (2) son premisas plausibles, (3) es una premisa problemática, ya que la posibilidad primaria no coincide con el tipo de posibilidad metafísica a la que podría apelar un materialista. Esta posibilidad metafísica, tal y como se define más arriba, coincidiría con la posibilidad secundaria. Ahora bien, parece que en el caso de F y C la relación entre las intensiones primarias y secundarias es especial. En el caso de C , no está claro que podamos separar las intensiones primarias y las secundarias, ya que en el caso de la experiencia consciente no parece existir una diferencia entre ser y parecer, como en el caso del agua. De todos modos, no es necesario que ambas intensiones coincidan en el caso de la experiencia consciente para que el argumento funcione. Lo importante para el argumento es que se muestre que, en los casos de estados mentales con propiedades fenoménicas, las intensiones primarias son las relevantes para evaluar los distintos mundos posibles:

Esto es precisamente lo que se ha hecho con la conciencia. Hemos visto que hay mundos que son físicamente idénticos al nuestro que carecen de conciencia, según la intensión primaria de la misma. Esta diferencia en los mundos es suficiente para mostrar que hay propiedades más allá de las propiedades físicas. Por analogía, si pudiéramos mostrar que hay mundos físicamente idénticos al nuestro en los que no

hay una sustancia acuosa, habríamos establecido el dualismo respecto al agua del mismo modo que si hubiéramos establecido que hay mundos físicamente idénticos al nuestro en los que no hay H₂O. Y de un modo importante, la diferencia respecto a la intensión primaria puede establecerse independientemente de factores *a posteriori*, de manera que las consideraciones respecto a la necesidad *a posteriori* son irrelevantes. (Chalmers 1996: 132-33)

El argumento de Kripke contra la tesis contingente de la identidad entre lo físico y lo mental se nutre, en última instancia, de las mismas intuiciones, a pesar de que Chalmers considera que se compromete con algunas tesis que no son necesarias para establecer la conclusión. Según Kripke, las identidades cuyos dos miembros son designadores rígidos son necesarias. Volviendo a (a), el agua no sólo es H₂O, sino que lo es necesariamente. Como ya se ha dicho, esta identidad puede parecer, en principio, contingente, ya que podríamos imaginar que el agua no es H₂O, sino XYZ; sin embargo, según Kripke, no estamos describiendo adecuadamente la posibilidad que imaginamos. Lo que de hecho imaginamos es que cierta sustancia que se parece al agua, y se comporta como tal, es XYZ. Del mismo modo, una identidad como la del dolor con, pongamos por caso, la actividad de las fibras C, tendría que ser igualmente necesaria. Sin embargo, ¿qué ocurre con la posibilidad de que haya un dolor sin estado cerebral, o actividad de las fibras C? El problema es que esta posibilidad, según Kripke, no puede ser redescrita del mismo modo que en el ejemplo del agua, ya que no puede haber una sustancia dolorosa que no sea dolor, del mismo modo que puede haber una sustancia acuosa que no sea agua. Sentirse como un dolor es ser un dolor, es algo que forma parte de su esencia, y la posibilidad imaginada en este caso no puede ser redescrita de otra manera.

Volviendo al marco de Chalmers, ésta es una razón adicional que aduce para justificar que en el caso de la experiencia consciente las intensiones primarias son las relevantes y que, incluso, en el caso de la experiencia consciente, las intensiones primarias y secundarias coinciden:

La irrelevancia de la necesidad *a posteriori* se puede apoyar, adicionalmente, con la observación de que, con la conciencia, las intensiones primaria y secundaria

coinciden. Lo que hace que un estado consciente sea una experiencia consciente en el mundo real es que tenga un sentir fenoménico. La diferencia entre las intensiones primarias y secundarias para el concepto de agua refleja el hecho de que podría haber algo que olierá y se sintiera como agua en algún mundo contrafáctico que, de hecho, no es agua, sino simplemente materia acuosa. Pero si algo se siente como una experiencia consciente, incluso en un mundo contrafáctico, entonces es experiencia consciente. Todo lo que significa ser una experiencia consciente, en cualquier mundo posible, es sentirse de cierta manera. (Chalmers 1996: 133)

Si las intensiones primarias y secundarias, tanto de F como de C, coincidieran, esto permitiría inferir que lo que es concebible primariamente respecto a la experiencia consciente también es concebible secundariamente, de manera que el argumento así completado, podría establecer la conclusión antimaterialista:

- (1) $F \ \& \ \sim C$ es concebible.
- (2) Si $F \ \& \ \sim C$ es concebible, $F \ \& \ \sim C$ es primariamente posible.
- (3) Si $F \ \& \ \sim C$ es primariamente posible, $F \ \& \ \sim C$ es secundariamente posible.
- (4) Si $F \ \& \ \sim C$ es secundariamente posible, el materialismo es falso.
- (5) Por lo tanto, el materialismo es falso.

Como ya adelanté, según Chalmers, no es necesario, para que el argumento funcione, que las intensiones primarias y secundarias de F y C coincidan, aunque, por otra parte, considera que esta tesis es bastante plausible para el caso de C (véase Chalmers 2010). En el caso de F, Chalmers contempla la posibilidad de que las intensiones primarias de las propiedades que ahora reconocemos como físicas se encuentren ligadas a ciertos roles teóricos de manera que, por ejemplo, la extensión de un término como “masa” sería cualquier propiedad que desempeñe el

rol de masa. Por otra parte, podría suponerse que la intensión secundaria de un término como masa coincide con la propiedad que, de hecho, desempeña el rol de masa, y habría que distinguir unas propiedades de la física que podríamos llamar estructurales, asociadas a un rol, y unas propiedades físicas intrínsecas. Si ambas intensiones difiriesen, las propiedades estructurales de la física serían compatibles con la posibilidad metafísica de los zombies y, en ese caso, solo las propiedades intrínsecas de la física junto con las propiedades estructurales implicarían, necesariamente, las propiedades fenoménicas. Esta última implicación, según Chalmers, coincide con el monismo de Russell, según el cual la experiencia consciente se encuentra ligada a las disposiciones físicas y a las propiedades intrínsecas que sirven como bases categóricas a dichas disposiciones (véase Chalmers 2010). La naturaleza de dichas propiedades podría ser protofenoménica o incluso fenoménica. Si dichas propiedades se aceptaran como físicas, el monismo de Russell podría considerarse una forma de fisicismo. No obstante, su carácter especulativo y su posible naturaleza podría acercar la posición de Russell al dualismo. Si las intensiones primarias y secundarias de F no coincidieran, una posición como la de Russell sería verdadera y, por ende, tampoco podríamos decir que el materialismo sería verdadero, a menos que se ampliara la concepción de materialismo en discusión. La conclusión del argumento, así formulado, sería que o el materialismo no es verdadero o un monismo tipo russelliano es verdadero.⁵⁴

Dado que el argumento, en cualquiera de sus formas, es válido, parece que una estrategia que se oponga al mismo debería cuestionar, al menos, alguna de sus premisas. Sin embargo, la posición que desarrollaré en los capítulos posteriores no consistirá en mostrar la falsedad de alguna de las premisas, sino que, más bien,

⁵⁴ En tal caso, el argumento que construye Chalmers (2010) es el siguiente:

- (1) F & ~C es concebible.
- (2) Si F & ~C es concebible, F & ~C es primariamente posible.
- (3) Si F & ~C es primariamente posible, F & ~C es secundariamente posible o el monismo russelliano es verdadero.
- (4) Si F & ~C es secundariamente posible, el materialismo es falso.
- (5) Por lo tanto, el materialismo es falso o el monismo russelliano es verdadero.

cuestionaré la inteligibilidad de las mismas. No argumentaré en contra del marco bidimensionalista esbozado ni de las relaciones que Chalmers describe entre concebibilidad y posibilidad metafísica, e incluso los argumentos que presentaré en los siguientes capítulos serán compatibles con la teoría del significado que propone Chalmers. La concepción de la experiencia consciente desarrollada por Chalmers es independiente del marco bidimensionalista y está muy lejos de ser un dato básico a tomar como punto de partida. Según esta concepción, la experiencia consciente constituye una dimensión especial del mundo en la medida en que no puede separarse su ser de su aparecer; en el ámbito de la experiencia consciente, no hay lugar para la ilusión ni, en principio, para el error. Además, una concepción que permite contemplar posibilidades como las imaginadas por Chalmers o Kripke sostiene un punto de vista muy fuertemente realista sobre los estos estados mentales, ya que nada impide concebir un mundo idéntico al nuestro, pero con algunos dolores desencarnados para los que, dada su condición, no existen unas circunstancias epistémicas, ni siquiera ideales, que nos permitan conocer su existencia. Más adelante llevaré a cabo un análisis de esta concepción tan realista de la experiencia consciente, así como de sus implicaciones. La idea que destaco en este momento es que las fuentes intuitivas de las que se nutre el problema de la conciencia, lejos de ser un fenómeno básico, constituyen una concepción del mundo y de la mente fuertemente cargada metafísicamente, que además es lógicamente independiente del aparato bidimensionalista esbozado.

Este diagnóstico tiene, al menos, dos consecuencias importantes en la dialéctica del debate sobre la experiencia consciente. En primer lugar, una posición como la de Chalmers no puede responder tan fácilmente a determinadas objeciones, alegando que el problema básico de la conciencia no ha sido rozado porque se ha dejado fuera el aspecto fenoménico, pretendiendo que éste es un dato básico innegable. Las posiciones eliminacionistas ⁵⁵ son las que más frecuentemente pueden recibir este tipo de respuesta y, en ese sentido, la posición que voy a desarrollar podría ser considerada eliminacionista. La cuestión es que habría que determinar qué es exactamente lo que éstas “eliminan” para que puedan

⁵⁵ Un ejemplo paradigmático de este caso es Daniel Dennett (1991).

ser llamadas eliminacionistas, ya que lo que de entrada hace implausible el eliminacionismo es la negación de la existencia de una parcela de la realidad perteneciente al sentido común, y a la que difícilmente podemos renunciar. Por el contrario, si lo que se pone en cuestión es un principio tan cargado teóricamente y discutible como el que acabamos de citar, el debate no puede quedar tan fácilmente cerrado. Asimismo, la afirmación de que la ciencia todavía no ha dado el menor paso en la comprensión de la experiencia consciente podría ponerse en cuestión en la medida en que no está claro que deje fuera una parcela de la realidad tan evidente como pretende Chalmers. Resulta paradójico que un proyecto con un claro cariz cientifista como el de Chalmers considere prácticamente inútil cualquier aportación científica que no provenga de la física, como inútil para la comprensión del *hard problem* de la conciencia.

En segundo lugar, el hecho de que no se haya dado más importancia a los supuestos metafísicos implicados en la formulación de la conciencia, y de que se hayan descartado apresuradamente posibles aportaciones de distintos campos, ha convertido gran parte del debate en una compleja reflexión librada, en su mayor parte, en el ámbito de la lógica y la teoría del significado. No se trata de negar la importancia de la reflexión abstracta en torno a la relación entre concebibilidad y modalidad, o el desarrollo de herramientas conceptuales que puedan ser posteriormente aplicables a diversos ámbitos; pero quizá habría que ser más consciente de en qué medida dichas discusiones son neutrales respecto a las cosas que de hecho existen y a sus cualidades.

Ned Block y Robert Stalnaker (1999) ofrecen un análisis a la respuesta desarrollada por Chalmers que, en cierta medida, coincide con la crítica aquí esbozada. Por una parte, el marco bidimensionalista podría, a lo sumo, representar las intensiones primarias en caso de que existieran, pero éstas requerirían de una justificación que Chalmers no da. Por otra parte, ambos autores señalan que no se pueden postular intensiones primarias por el hecho de que haya convergencia en nuestras evaluaciones modales, ya que este tipo de razonamientos no sólo depende del uso de conceptos, sino que se encuentra influido por prioridades epistémicas, compromisos, teorías y principios metodológicos que incluso podemos haber

absorbido inadvertidamente con nuestra cultura científica.⁵⁶ Esta idea, a su vez, hace difícil la inteligibilidad de un razonador ideal. La influencia de estos factores es una posibilidad que el puro razonamiento a priori no puede excluir, y una de las tesis básicas que defenderé es que, precisamente, hay una concepción metafísica determinada de la experiencia consciente que configura la comprensión y la evaluación de los experimentos mentales utilizados en los argumentos presentados, y trataré de mostrar que dicha concepción metafísica es incoherente, al igual que los experimentos mentales que se encuentran detrás de la formulación del problema de la conciencia. Si se logra mostrar, de un modo positivo, que la concepción de la experiencia consciente que subyace al argumento de Chalmers es incoherente, el argumento de Chalmers será irrelevante para determinar lo que es posible respecto a la conciencia. No obstante, el resultado de esta crítica no será ni la reivindicación de una posición fisicista, ni su negación. Por el contrario, supondrá una crítica general al marco en el que tienen lugar las discusiones en filosofía de la mente y, especialmente, en relación a la experiencia consciente, de manera que propondré que los intentos de comprender lo mental no tengan como objetivo, necesariamente, encontrar un espacio para lo mental en un mundo fisicista.

⁵⁶ Igualmente, las reflexiones de Kripke son muy valiosas como contribución a la comprensión del significado de nuestros términos de género natural, pero difícilmente puede establecer la verdad de un enunciado con un contenido tan concreto como que el agua es necesariamente H₂O. ¿De qué manera podría rechazar, por ejemplo, la aparición futura de una teoría atómica más básica que reformule nuestra concepción de las moléculas, y que distinga dos clases de H₂O, de modo que llegue a saberse que solo uno de los dos es el que se encuentra en la Tierra?

4. Epifenomenismo y experiencia consciente

En el capítulo anterior presenté los argumentos que tratan de mostrar la irreductibilidad de las propiedades fenoménicas de la experiencia y, como consecuencia, la falsedad del fisicismo. Ya vimos que, con independencia del debate respecto a la verdad o falsedad del fisicismo, la concepción de la experiencia consciente y de lo mental defendida por Chalmers tiene consecuencias tanto respecto a nuestras concepciones de sentido común, como a las prácticas vinculadas a éstas. Una de las consecuencias más llamativas de la posición de Chalmers, de la que no he hablado hasta el momento, es que la aceptación conjunta de la irreductibilidad de la experiencia consciente y de la idea de que el resto de lo mental puede someterse exitosamente, al menos en principio, a un análisis funcional parece que implica la total ausencia de relevancia causal de la experiencia consciente. Para ilustrar esta consecuencia podemos recurrir, nuevamente, a la fantasía de los zombis. Un zombi gemelo de A⁵⁷ actuará del mismo modo que A, realizando al detalle sus mismos movimientos, e incluso profiriendo las mismas palabras⁵⁸. Ahora bien, por hipótesis, si conociéramos todos los hechos físicos vinculados a A y dispusiéramos de una física lo suficientemente potente, podríamos predecir al detalle todos los movimientos de A, sin necesidad de tener en cuenta nada referente a las propiedades fenoménicas de su experiencia consciente. Si esto es así, la experiencia consciente no parece desempeñar ningún rol causal y no está muy claro en qué sentido podría tener relevancia explicativa. Esta afirmación es lo que en la reflexión filosófica sobre lo mental se conoce como epifenomenismo.

En general, los distintos dualismos respecto a lo mental han tenido que posicionarse, en un sentido u otro, respecto al epifenomenismo. En algunos casos, ha sido considerado como una consecuencia lo suficientemente implausible, bien para rechazar el dualismo de partida del que se deriva, o bien para calificar dicho dualismo de algún modo que permita evitarlo como conclusión, como en el caso del dualismo interaccionista.

⁵⁷Esto es, un duplicado físico y funcional de A que, no obstante, carece de experiencia fenoménica.

⁵⁸En el caso de los zombis no es muy claro si podría decirse que profieren palabras, o simplemente generan sonidos que en sus propiedades físicas son indistinguibles de las palabras proferidas por un ser humano.

Según Chalmers, si bien el epifenomenismo es una consecuencia bastante contraintuitiva de su teoría, no se ha mostrado que constituya una objeción fatal a la misma, ya que en ningún caso hay argumentos que muestren claramente su incoherencia. En este capítulo, trataré de determinar, en primer lugar, el vínculo general de la posición de Chalmers con el epifenomenismo, comparándolo con otras discusiones en filosofía de la mente en las que aparece dicho problema; en segundo lugar, analizaré algunos de los problemas concretos a los que Chalmers se tiene que enfrentar y sus respuestas, prestando especial atención a lo que él mismo ha bautizado como la paradoja del juicio fenoménico; y, finalmente, extraeré algunas conclusiones respecto a la concepción de lo mental de Chalmers, mostrando el conflicto entre ésta y algunas condiciones necesarias para que sea posible atribuir experiencia consciente a otros seres.

4.1. El epifenomenismo

Aunque la formulación del problema del epifenomenismo⁵⁹ en lo mental se remonta a una larga tradición, pueden distinguirse varios problemas que responden a dicha denominación y que se vinculan a distintas inquietudes filosóficas. En el caso de Descartes, o bien de cierta concepción heredada de la filosofía cartesiana, el epifenomenismo es una conclusión a evitar a toda costa, en la medida en que pone en peligro la eficacia causal de lo que constituye la esencia del hombre, a saber, una *res cogitans*. El dualismo interaccionista cartesiano constituye una respuesta motivada por el problema particular de determinar una esencia del hombre que lo haga distinto de lo demás y del resto de seres vivos, así como de salvaguardar la libertad humana, haciéndola posible en un mundo que se encuentra regido por procesos y propiedades ajenos a lo mental.⁶⁰ Por el contrario, el modo específico que tiene de irrumpir el epifenomenismo en la filosofía de la mente actual se

⁵⁹Si bien el problema es de largo recorrido, su denominación es algo más reciente; concretamente, se trata de un término acuñado en el siglo XIX por William James, usado para referirse a síntomas secundarios de una enfermedad, que no juega ningún papel como causante de la misma.

⁶⁰Para algunos desarrollos posteriores del cartesianismo que se comprometen con algún tipo de armonía preestablecida será muy difícil dejar lugar para algún tipo de libertad.

encuentra fuertemente vinculado a la idea de la clausura causal del mundo físico, de manera que el objetivo principal de la filosofía ya no es salvaguardar determinada concepción del ser humano, sino tratar de buscar un lugar para lo mental en la medida en que el marco fisicista establecido nos deje margen de maniobra. En caso de que esto no sea posible, la prioridad teórica de mantener la autonomía de lo físico sólo deja lugar a una aceptación, sin más, de lo que no puede ser reducido y, a lo sumo, resta el movimiento de hacer aceptable, en la medida de lo posible, dicho residuo. El epifenomenismo actual es una consecuencia de tratar de tomar en serio el ámbito de lo mental sin cuestionar, a su vez, la validez del fisicismo. Parece claro que lo mental es eficaz causalmente en un sentido específico no aplicable, por ejemplo, a la cadena causal en la que se inserta el movimiento de una piedra movida por un desprendimiento. Si alguien cree que va a llover y sale a la calle con un paraguas, el hecho de que use el paraguas está causado por la creencia de que está lloviendo (independientemente de que llueva o no), y ello en virtud de ciertas propiedades semánticas específicas. Ahora bien, ¿en qué medida pueden llegar a ser estas propiedades específicas relevantes causalmente una vez se acepta el marco fisicista?

Ya hemos visto que, de acuerdo con la distinción de Chalmers entre el concepto psicológico de lo mental y el concepto fenoménico, los estados intencionales como las creencias, dejando de lado todo el trabajo analítico y empírico que todavía quedaría para determinar completamente de qué manera se realizan funcionalmente, habrían entrado en un camino seguro para ser reducidos y comprendidos.⁶¹ Sin embargo, también señalamos que este proyecto no estaba exento de profundos problemas. El tipo de solución que propone el funcionalismo, aunque aparentemente hace justicia a la idea de que las propiedades mentales son eficaces en virtud de cierta especificidad, no logra rescatar exitosamente la idea general de que lo mental sea eficazmente causal. Si bien no es éste el problema específico que plantea la dimensión fenoménica de lo mental, tal y como Chalmers la acota, su formulación no es independiente de la dialéctica en torno al

⁶¹ A lo largo del capítulo vamos a ver un tipo de estados intencionales que tienen como objeto contenidos que incluyen propiedades fenoménicas.

funcionalismo que genera el problema de la conciencia. A continuación, seguiré el hilo discursivo que conduce desde el proyecto reduccionista del materialismo, en su versión funcionalista vinculada a la noción de supervinencia, estableciendo el vínculo conceptual de este proyecto con la conclusión de la irreductibilidad de la conciencia y su carácter epifenoménico.

4.1.1. El fisicismo y la eficacia causal de lo mental

Si recordamos el argumento causal a favor del fisicismo, éste necesitaba de tres premisas para poder establecer su conclusión. Su estructura es la siguiente⁶²:

- (1) Los estados mentales tienen efectos físicos.
- (2) Todo efecto físico tiene como causa⁶³ antecedente un estado físico.
- (3) Los efectos físicos de los estados mentales no están sobredeterminados por causas distintas.
- (4) Los estados mentales son idénticos a estados físicos.

La justificación de (1) apela a los casos intuitivos en los que un estado mental causa un determinado efecto físico, como cuando alguien tiene sed y dicho estado mental causa la serie de movimientos conducentes a beber un vaso de agua.⁶⁴

Por otra parte, (2) forma parte del contenido mínimo del fisicismo y se justifica, en principio, por la evidencia empírica de la que disponemos. En este

⁶² No hay un acuerdo respecto a la formulación adecuada del argumento causal. En este caso me he servido de la formulación de David Papineau por su generalidad, aunque he formulado el argumento aplicado a estados mentales en general, no sólo a estados mentales conscientes.

⁶³ Quizá no sea suficiente afirmar que todo efecto físico tiene como causa un estado físico, sino que esta causa ha de ser completa (véase Corbí y Prades 2000).

⁶⁴ Podría decirse que esta premisa ya da por sentado que el epifenomenismo es falso. No obstante, la aceptaremos por mor del argumento. Una reformulación más concreta del argumento combinada con otros supuestos nos llevará a un tipo de epifenomenismo que no se aplica específicamente a la fenomenología, pero que conduce a la concepción de Chalmers de la experiencia consciente.

sentido, David Papineau sostiene que algunos de los descubrimientos del siglo XX vinculados a la comprensión más detallada de la biología han reforzado notablemente la justificación de esta premisa; por otra parte, como ya vimos en el Capítulo 1, el contenido de un enunciado de este tipo no está determinado de un modo muy concreto, en la medida en que no hay una definición muy específica de físico, aunque bastará con que aceptemos el contraste básico entre lo físico y lo no físico, o mental, como punto de partida.

La introducción de (3) excluye la posibilidad de que dos líneas causales converjan en un mismo efecto. En tal caso, podría hablarse de una sobredeterminación causal. Las razones para rechazar este tipo de posibilidad en el caso de lo mental son dos. En primer lugar, aunque se reconoce la existencia de fenómenos de sobredeterminación causal (pensemos, por ejemplo, en el impacto de varias balas que, independientemente producirían una muerte) éstos no abundan en la naturaleza; sin embargo, la negación de (3) convertiría cualquier suceso en el que interviniera lo mental en un caso de sobredeterminación. En segundo lugar, si atendemos al tipo de explicaciones causales vinculadas a lo mental, parece que la sobredeterminación no es intuitiva. En el caso de los impactos de bala, es evidente que, si no se hubiera producido el impacto de una de las balas, la otra, u otras, habrían producido la muerte de la persona. Por el contrario, si pensamos en el caso del individuo que bebe un vaso de agua porque tiene sed, aceptamos que en los movimientos que lleva a cabo con este fin hay implicado un determinado proceso neuronal, y un contrafáctico que afirme que la persona hubiera bebido el vaso de agua, aun cuando no se hubiera producido dicho proceso neuronal, resulta contraintuitivo.

El núcleo del argumento es que, al cerrarse la posibilidad de una doble línea causal, la única alternativa viable es la identificación de estados físicos y estados mentales, de manera que puede reconocerse la eficacia causal de estos últimos sin violar la clausura causal del mundo físico. Ahora bien, hablando en general de este tipo de identidad, ¿se ha conseguido preservar un sentido lo suficientemente robusto de lo que significa reconocer la eficacia causal de lo mental? El problema es que considerar que lo mental es eficaz como causante de un efecto incluye, de

un modo general, la idea de que es eficaz en virtud de ser mental, del mismo modo que un coche llega a alcanzar una velocidad determinada en virtud del número y tipo de cilindros de su motor, y no en virtud de su color. Otro aspecto de esta cuestión es en qué medida lo mental puede constituirse en un objeto de estudio autónomo respecto a ciencias como la física o la biología.⁶⁵

En principio, el modelo funcionalista de reducción proporciona un marco independiente que permite centrar la investigación científica en unas propiedades específicas, que no coinciden con las propiedades de la física o de otras ciencias. Daniel Dennett (2001), por ejemplo, señala que el intento de comprender la mente desde la física se encuentra anclado a una concepción anticuada de lo mental. Parece que el funcionamiento de la mente depende de cierta organización o estructura que puede implementarse en sistemas muy dispares en cuanto a su disposición; en este sentido, determinar lo que es común a esta variedad de sistemas, de un modo relevante para el estudio de la mente, es una tarea que resulta poco prometedora si el punto de partida es, exclusivamente, la física.

Las distintas posiciones materialistas que han configurado el debate en filosofía de la mente en las últimas décadas han tenido que enfrentarse de una u otra forma a esta cuestión. En el caso de las teorías de la identidad de tipos, el problema era que, precisamente, parecía negarse la posibilidad de que lo mental fuera realizable de múltiples formas. Si un estado mental es idéntico a un determinado tipo físico, ya no se puede realizar en distintos sistemas físicos. La maniobra de trasladar la relación de identidad a los particulares (*tokens*), en lugar de los tipos, hace que se pierda, precisamente, la idea de que los estados mentales producen los efectos que producen en virtud de ser los estados mentales que son.⁶⁶

⁶⁵ En ocasiones, la tendencia a pensar en el conjunto de las ciencias como una estructura unitaria y jerárquica compuesta por compartimentos estancos se distancia mucho de la práctica real de la ciencia, incluso en ciencias como la biología, en la que ramas como la zoología difícilmente pueden dejar de tener en cuenta estados de diversos tipos, y especialmente, estados intencionales.

⁶⁶ La teoría más conocida de este tipo es la expuesta por Donald Davidson (1980) y conocida como monismo anómalo. De acuerdo con esta teoría, dos sucesos particulares s_1 y s_2 , se encuentran vinculados causalmente en la medida en que puedan subsumirse mediante una ley estricta; ahora bien, las únicas leyes estrictas son las de la física, de manera que s_1 es la causa de s_2 en la medida en que ambos puedan ser descritos en términos físicos y su vínculo pueda subsumirse de acuerdo con una ley estricta de la física. Por supuesto, dichos sucesos también podrían ser descritos en términos mentales, pero puesto que no hay leyes psicofísicas que vinculen eventos descritos en términos mentales, toda la eficacia causal se restringiría a las propiedades físicas de los

Ahora bien, ¿en qué medida puede el modelo funcionalista hacer compatibles tanto la tesis fisicista como la eficacia causal de lo mental? A continuación, presentaré una de las dificultades más importantes a las que ha tenido que enfrentarse el funcionalismo y que apunta a que tampoco el fisicismo funcionalista es capaz de salvar lo mental como una fuente genuina de causalidad.

4.1.2. Supervinencia, funcionalismo y causalidad

De un modo superficial, ya presenté algunas posibilidades de objeción al funcionalismo como una teoría demasiado débil, en la medida en que resulta intuitivo que duplicados funcionales de un determinado individuo podrían no compartir sus estados mentales. Block (1980) presenta argumentos que tratan de mostrar de un modo ilustrativo este punto. Uno de ellos es el experimento mental de la nación china reproduciendo la estructura funcional de un cerebro humano, y otro es el espectro invertido (ambos argumentos fueron presentados en el capítulo anterior). Los dos argumentos parecen plantear la limitación del funcionalismo para dar cuenta de las propiedades fenoménicas de la experiencia. De hecho, Chalmers considera el argumento del duplicado funcional de la mente en una nación como una variante del experimento de los zombis. Sin embargo, es posible que estos argumentos puedan aplicarse, en general, al ámbito de lo mental. John Searle (1992) considera que el planteamiento dominante de la filosofía de la mente contemporánea está completamente equivocado. El problema básico que plantea es la insuficiencia de las teorías de la mente dominantes, no sólo para dar cuenta de la experiencia consciente, sino también de los estados intencionales. La intuición básica de Searle, en su forma más conocida, se expresa en el argumento de la habitación china. En dicho argumento, se invita a imaginar a un sujeto encerrado en una habitación que dispone de un manual (quizá infinito) para manipular caracteres chinos. Dicho sujeto recibiría preguntas, u otro tipo de información, redactadas en chino y, a través de la consulta del manual, respondería de un modo aparentemente

sucesos. Como consecuencia, esta teoría no es compatible con la idea de que un suceso cause otro en virtud de sus propiedades mentales.

inteligente a dichas preguntas. La conclusión que trata de establecer el argumento es que este pintoresco sistema no puede considerarse un sistema pensante, capaz de entender el lenguaje y dotado de intencionalidad. El argumento es una propuesta de refutación de la Inteligencia Artificial fuerte⁶⁷, pero su lema podría resumirse en la idea de que la mera manipulación sintáctica de símbolos (y del mismo modo los programas ejecutados por los ordenadores actuales) no implica, por sí sola, la comprensión de significados ni la semántica. Según Searle, la ciencia cognitiva y la filosofía de la mente contemporánea no podrían dar cuenta de la especificidad de lo mental. Una de las versiones del funcionalismo, de hecho, la versión fundacional (véase Putnam 1960), es de carácter computacional. Puesto que la computación, en última instancia, se ocupa de relaciones sintácticas y de su manipulación, en sí misma no puede dar cuenta de la idea de que algo sea eficaz causalmente en virtud de su contenido semántico. Esto es lo que Ned Block (1990) denomina la Paradoja de la Eficacia Causal del Contenido. Si pensamos en un ordenador actual, su funcionamiento se basa en la idea de que no importa cuál sea el material con el que se haya fabricado, siempre y cuando implemente determinado tipo de relaciones sintácticas. Por supuesto, es más fácil construir un sistema más eficiente, tanto por su rapidez como por su capacidad y tamaño, aprovechando las posibilidades de un *hardware* basado en la electricidad, que construyendo un sistema hidráulico o de poleas; pero esto sólo supondría, en principio, una limitación técnica irrelevante. La cuestión es, que para que el ordenador funcione, todo lo que se necesita es que, dados ciertos *inputs*, en determinados puntos de su estructura y de una manera compleja, haya ciertos puntos en los que se deje o no se deje pasar la corriente eléctrica. La cuestión es la siguiente: si este modelo se puede extender para comprender el funcionamiento de la mente, ¿qué tiene que ver todo esto con el contenido semántico del pensamiento teniendo en cuenta que el funcionamiento adecuado del ordenador podría explicarse exclusivamente por leyes físicas? Y en caso de que el modelo funcione y sea adecuado, ¿no deberíamos aceptar que la

⁶⁷ Por Inteligencia Artificial fuerte se entiende el intento de llegar a producir un artefacto con inteligencia y pensamiento genuinos, indistinguible, en dicho sentido, de un ser humano.

eficacia causal de lo mental es una ilusión, de manera que lo único relevante son los poderes causales del sistema físico que realiza el ordenador?

Hasta ahora, para reproducir esta problemática me he ceñido a la concepción computacional de la mente. No obstante, sin tener en cuenta la formulación concreta del modelo computacional, este tipo de problema se puede generar a un nivel mayor de abstracción, de manera que podría reproducirse en todo intento funcionalista de conciliar la clausura causal del mundo físico con la eficacia causal de lo mental. El problema general es que, para poder lograr este objetivo, las propiedades funcionales dependen crucialmente de sus realizaciones físicas, de manera que dichas propiedades no tienen unos poderes causales autónomos. Si mi sensación de sed produce un movimiento en mi brazo consistente en acercar un vaso de agua a mi boca, dado el modelo funcionalista habrá un realizador físico de mi sed que cumpla con el rol causal adecuado de la sed, y parte de los poderes causales de dicho realizador incluirán el movimiento de mi brazo para acercar el vaso de agua a mi boca, de manera que, como en el caso de la teoría de la identidad de casos, no se puede dar contenido a la idea de que un estado es causalmente eficiente en virtud de sus propiedades mentales. Este argumento ha sido nombrado por Jaegwon Kim (1998) como el argumento de la exclusión causal.

Ahora bien, hasta ahora hemos confiado, sin cuestionarlo, en que está claro qué significa que los realizadores físicos sean los que realmente gozan de eficacia causal, o que puede darse por sentado, sin más, que entendemos qué significa que un estado sea eficaz causalmente en virtud de sus propiedades físicas. Como ya vimos en el Capítulo 1, en el debate general en torno al fisicismo no se encuentra una caracterización precisa de a qué nos referimos con físico y no parece que dicho debate se nutra especialmente de la información empírica que actualmente está proporcionando la física. Pero dejando de lado este problema, habría que decir algo respecto a cómo es posible que las propiedades físicas puedan ser eficaces causalmente, en contraste con las propiedades mentales. Por otra parte, parece aceptarse con bastante unanimidad que el modelo funcionalista es aplicable a otro tipo de propiedades que pertenecen a distintas ciencias, como puede ser la propiedad de ser un organismo vivo, o la instanciación de un proceso geológico; en

tal caso, ¿qué ocurre con estas propiedades? ¿También habría que decir que no gozan de eficacia causal?

4.1.3. La concepción estratificada de la realidad y el drenaje de los poderes causales

Un argumento en la línea de las cuestiones mentadas en las últimas líneas es presentado por Kim (1998) como una posible objeción al argumento de la exclusión causal. Según dicho argumento, una consecuencia que puede extraerse del argumento de la exclusión causal es su generalización, de manera que ejemplos paradigmáticos de causalidad, como la relación entre el consumo de tabaco y el cáncer, o incluso la misma causalidad, podrían llegar a cuestionarse:

[S]e seguiría que toda la causalidad de un macronivel sería una mera ilusión. ¡La bola de béisbol perdida no rompió la ventana, después de todo, y el terremoto no hizo que los edificios colapsaran! Esto nos parece intolerable. Es más: ¿y si no hay un nivel último (bottom level) (como Block nos reta a considerar)? Parece como si no hubiera un nivel último en esta imagen, ¡los poderes causales se drenarían por un agujero sin fondo y no habría causalidad *en ninguna parte!* (Kim 1998: 81)

En esta sección, no trataré de ofrecer una discusión detallada del argumento ni de sus posibles soluciones, sino de cómo se genera y de qué manera se relaciona, tanto con la concepción general de la mente de Chalmers, como con su formulación del problema de la conciencia.

Según Block (2003), que el drenaje de los poderes causales se siga del argumento de la exclusión es una consecuencia lo suficientemente implausible como para rechazar el argumento. Tal y como interpreta Block el argumento, parece que la única manera de frenar la generalización del argumento es que exista un nivel último o fundamental (*bottom level*):

Si no hay un nivel fundamental, y si cualquier propiedad (supuestamente) causalmente eficaz es superviniente sobre una propiedad de nivel “más bajo”

(Llamemos a esto: “subvinencia sin fin”), entonces (puede decirse) el Argumento de la Exclusión Causal de Kim mostraría, si es válido, que cualquier afirmación de la eficacia causal de propiedades se ve socavada por la afirmación de un nivel más bajo y, de este modo, no hay causalidad. (Block 2003: 5)

De acuerdo con el análisis de Block, lo que convierte al argumento de la exclusión en implausible es que el bloqueo de una consecuencia tan inaceptable como la negación de la causalidad dependa de una cuestión abierta en física, como lo es la posible determinación de un nivel fundamental, y continúa:

Esto no es una prueba, pero asumiéndola, no obstante, por el momento, llegamos a la siguiente refutación. Que haya o no un nivel fundamental es una cuestión abierta, pero no es una cuestión abierta si hay causalidad. Puede ser una cuestión abierta si fumar tabaco causa cáncer, pero no es una cuestión abierta que cualquier cosa cause algo alguna vez. (Block 2003: 5)

Sin embargo, por muy implausible que pueda resultar el argumento de la exclusión, hay algo extraño en el tipo de refutación de la que se sirve Block. El tipo de problema en el que piensa Block cuando habla de una cuestión abierta se contrapone a cuestiones escépticas como la duda respecto a si el mundo existe sólo desde hace cinco minutos. No obstante, para apoyar la idea de que la existencia de un nivel fundamental de la física constituye una cuestión abierta, Block recurre a algunas hipótesis científicas y asume que hay un vínculo especial entre “cuestiones abiertas” y “cuestiones científicas”. Sin embargo, Block no profundiza más en qué pueda consistir la existencia de un nivel fundamental de la física y considera esto, sin más, como una cuestión que deberá determinar la física. No obstante, ¿qué evidencia empírica podría determinar que existe un nivel último o fundamental de la física? Schaffer (2003) considera, por ejemplo, que no puede concluirse por razones meramente lógicas o conceptuales que tiene que haber un nivel último de la realidad y reconoce el argumento empírico y explicativo como el mejor candidato para concluir su existencia. Schaffer concluye que no hay un buen apoyo empírico para la idea de que existe un nivel último; sin embargo, es más interesante su conclusión

de que una tesis fisicista exige la existencia de un nivel fundamental, de manera que el marco fisicista convierte la existencia de un nivel último en algo más que en una cuestión abierta a la espera de ser respondida por la investigación empírica. Y si no hay un nivel último, entonces el fisicismo es falso.⁶⁸

Como señalaba al final de la sección precedente, si el argumento de la exclusión funciona, parece que es necesario establecer algún tipo de diferencia metafísica entre las propiedades que han quedado vaciadas de poderes causales y las propiedades que gozan de poderes causales genuinos, en este caso, las propiedades físicas. De otro modo, si todas las propiedades existentes son propiedades funcionales, ¿qué puede impedir la generalización del argumento? ¿No podría decirse que es más bien la idea de que debe haber un nivel fundamental de la física la que determina, en cierta manera, el tipo de entidades que se encontrarán? En el caso de la mente, si esta historia llegara a tener sentido, la compleción del proyecto de comprensión de la eficacia causal de la mente consistiría en dar con los realizadores últimos que se encuentran en la base sustentadora de toda la ciencia. El problema es, ¿cómo podríamos saber que hemos dado con ellos?⁶⁹

La existencia de un realizador último en los términos planteados apunta a una discusión metafísica, más que a una cuestión susceptible de ser dilucidada en el seno de la física. Una de las intuiciones más atractivas de las que se nutre el fisicismo es que todo está compuesto por partículas físicas, de manera que el comportamiento de todo objeto se reduce a la interacción de las distintas partículas que la física describe. Ahora bien, si la cuestión de dar con un realizador último fuese científica, sería muy difícil llegar a saber si se ha dado con el mismo. Esto implicaría que ya no es posible encontrar propiedades más básicas o que, de algún modo, se ha dado con un objeto tal que no puede haber nada más simple que el mismo o cuyo comportamiento no puede ser entendido a partir de entidades simples. No está claro qué evidencia podría llevar a este tipo de conclusión. Esta

⁶⁸ Respecto a las relaciones entre el fisicismo y una concepción jerárquica de la realidad, véase el Capítulo 1.

⁶⁹ En este punto, no podemos evitar recordar la parte de la Dialéctica trascendental en la *Crítica de la Razón Pura* donde Kant explica como la propia actividad de la razón genera, *per se*, una ilusión trascendental que se muestra en la búsqueda de lo incondicionado.

discusión tiene ciertas semejanzas con los problemas planteados por los atomistas griegos, quienes postularon entidades indivisibles a partir de premisas metafísicas que poco tenían que ver con la observación. Del mismo modo, la discusión de Block no requiere demasiada evidencia científica ni asume un conocimiento detallado de la física actual.

La existencia de un nivel último de propiedades implicaría que hay unas propiedades que son muy diferentes a las propiedades funcionales e, incluso, muy diferentes respecto a propiedades poco problemáticas dentro del marco fisicista que, además, gozan de estatus científico. De algún modo, parece como si los realizadores últimos fueran la sustancia de la realidad. En el capítulo anterior, al presentar el argumento bidimensionalista contra el materialismo, hablé de la posibilidad de que las propiedades físicas sean propiedades disposicionales que satisfarían un rol teórico, aunque desconozcamos las propiedades intrínsecas que sirven como su base categórica. Esta es una distinción metafísica que no puede simplemente aceptarse como una intuición o como un punto de partida indiscutible. En la siguiente sección, discutiré la concepción de las propiedades físicas que tiene Chalmers y su relevancia para el problema de la conciencia.

4.1.4. Propiedades intrínsecas

Las cuestiones discutidas en este capítulo no reciben un tratamiento detallado en la filosofía de Chalmers, ya que en ella se da por supuesto que el funcionalismo es una teoría exitosa de toda dimensión de lo mental que no sea fenoménica. Sin embargo, algunos de estos problemas tienen importantes consecuencias en relación al problema de la conciencia, de modo que, de alguna manera u otra, Chalmers tiene que poder ofrecer algún tipo de respuesta a estas cuestiones, sobre todo teniendo en cuenta que su teoría de la conciencia pretende diseñar una hoja de ruta para la futura investigación científica sobre la conciencia. Chalmers se muestra cauteloso respecto a la cuestión del epifenomenismo en la medida en que es una tesis vinculada al problema de la causalidad, cuyo terreno es intrincado y del cual no se tiene una buena comprensión. De hecho, Chalmers considera que los

hechos respecto a la causalidad podrían no supervenir sobre el conjunto de los hechos físicos y, en este sentido, quizá haya algún tipo de nexo entre la causalidad y la experiencia consciente.⁷⁰

Es habitual encontrar en la filosofía de Chalmers posiciones metafísicas muy definidas y con implicaciones decisivas en el debate, en la discusión de cuestiones, en principio, periféricas. En este sentido, en una parte especulativa de *The Conscious Mind* (1996) encontramos una posición metafísica muy fuerte en relación a la cuestión del estatus de las propiedades físicas y la incompletud teórica a la que éstas pueden llevar:

La teoría física sólo caracteriza sus entidades básicas *relacionalmente*, en términos de sus relaciones causales o con otras entidades. Las partículas básicas, por ejemplo, se caracterizan en gran parte por su tendencia a interactuar con otras partículas. Su masa y su carga está especificada, por supuesto, pero toda esa especificación se reduce, en última instancia, a una tendencia a acelerarse de cierta manera por fuerzas, etcétera. Cada entidad se caracteriza por su relación con otras entidades, y estas entidades se caracterizan por su relación con otras entidades, y así indefinidamente (excepto, quizás, respecto a algunas entidades que se caracterizan por su relación con un observador). La imagen del mundo físico producida por esto es la de un mundo como un flujo causal gigante, pero no nos dice nada en absoluto respecto a qué *relaciona* esta causalidad. La referencia al protón se fija como la cosa que causa interacciones de cierto tipo, que se combina de determinadas maneras con otras entidades, etcétera; pero, ¿cuál es la cosa que está causando y combinándose? [...]

Alguien podría estar atraído por la visión del mundo como un flujo causal puro, sin propiedades adicionales que relacione la causalidad, pero esto llevaría a una visión del mundo extrañamente insubstancial. Todo lo que tendría en él son relaciones causales y nómicas entre entidades que están, en última instancia, vacías,

⁷⁰ “Hay algo irreductible en la existencia de leyes y causalidad. He evitado estos problemas en otro lugar incluyendo leyes físicas en la base de supervinencia, pero esto deja de lado puzle metafísico, en lugar de responderlo. [...] Aparte de la experiencia consciente, y quizá de la indexicalidad, estos constituyen los únicos hechos adicionales tales que tenemos razones para creer en ellos. No es extraño especular que estas dos clases de hechos no supervinientes, conciencia y causalidad, pueden tener una relación metafísica cercana” (Chalmers 1996: 78).

ocupantes inexistentes carentes de propiedades en sí mismos. Intuitivamente, es más razonable suponer que las entidades básicas que toda esta causalidad relaciona tienen alguna naturaleza interna propia, algunas propiedades *intrínsecas*, de modo que el mundo tiene algo de substancia. Pero la física puede, en el mejor de los casos, fijar la referencia de dichas propiedades en virtud de sus relaciones extrínsecas; no nos dice nada de lo que esas propiedades podrían ser. Tenemos algunas intuiciones vagas respecto a esas propiedades que se basan en nuestra experiencia de sus contrapartes macroscópicas—intuiciones respecto a la “masividad” de su masa, por ejemplo—pero es difícil dar cuerpo a estas intuiciones, y no está claro, reflexionando, que haya algo detrás de las mismas.

Sólo hay una clase de propiedad intrínseca, no relacional con la que tenemos una familiaridad directa, y esta es la clase de las propiedades fenoménicas. [...] (Chalmers 1996: 153)

Hay varios aspectos a destacar en este fragmento. En primer lugar, su situación en el conjunto del libro es extraña; este texto consta de algunas consideraciones especulativas que pueden ser el esbozo de una línea de respuesta a la objeción de que la posición de Chalmers conduce al epifenomenismo de la conciencia. Sin embargo, en él se describen algunos rasgos sustantivos de la ciencia física que tienen consecuencias importantes, no sólo respecto a la cuestión de la experiencia consciente, sino también en relación con el materialismo y su contenido. Si las propiedades físicas son esencialmente relacionales, ¿no constituye esto un principio importante que califica notablemente la tesis fisicista? Recordemos que, de acuerdo con el fisicismo, en un sentido mínimo, las propiedades físicas del mundo fijan el resto de propiedades; son, de alguna manera, las propiedades básicas del mundo. Sin embargo, estas reflexiones podrían incluso llegar a ser incompatibles con el fisicismo, en la medida en que el carácter relacional de las propiedades físicas parece incompatible con el tipo de asimetría requerido entre las propiedades físicas y el resto de propiedades.⁷¹ No obstante, en la reflexión principal sobre el fisicismo y sus tesis centrales no aparece ninguna idea semejante.

⁷¹ Un poco más abajo, Chalmers desarrolla algunas ideas que vinculan el carácter intrínseco de la experiencia consciente con los posibles componentes últimos de la materia, señalando a un posible giro conceptual en la reflexión sobre la relevancia causal; más adelante diremos algo respecto a esa cuestión.

Chalmers presenta sus argumentos como partiendo de tesis muy básicas y difíciles de rechazar, pero si posteriormente éstas son puestas en cuestión, se sirve de tesis fuertemente metafísicas para responder, de manera que no está claro qué es lógicamente anterior, si las intuiciones de las que parte en su presentación de los problemas, o dichas tesis.

Por otra parte, en el texto se ofrecen algunos argumentos que tratan de establecer el compromiso metafísico con la existencia de propiedades intrínsecas. Ahora bien, el estatuto de estos argumentos no está nada claro; de hecho, en ocasiones parecen presentarse como vagas intuiciones, o como mero sentido común, y otras veces parece que los argumentos son de tipo lógico. Respecto a la tesis de que todo en el mundo sea un puro flujo causal, signifique lo que signifique esto, Chalmers considera que “esto llevaría a una visión del mundo físico extrañamente insustancial”, y respecto a la tesis contraria, según la cual la causalidad debería relacionar “ocupantes” (*placeholders*) con unas propiedades intrínsecas propias, considera que, intuitivamente, es algo que “es más razonable suponer”. Estas expresiones tentativas, que apuntan a lo que es razonable o intuitivo, contrastan con otro tipo de expresiones que parecen señalar, con necesidad lógica, la falsedad de la tesis del flujo causal: por ejemplo, el hecho de que esta teoría no diga nada de aquellas entidades entre las que se produce la causalidad o la mención de entidades vacuas y de “ocupantes no existentes” (*nonexistent placeholders*) parece apuntar a un defecto lógico grave. De un modo mucho más contundente, más abajo se dice lo siguiente respecto a la posibilidad de un mundo de flujo puramente causal: “dicho mundo es, podría decirse, lógicamente imposible, ya que no hay nada en dicho mundo que la causalidad relacione” (Chalmers 1996: 153). No obstante, ¿por qué una teoría que afirme que sólo existe un flujo puramente causal tendría que aceptar dicha objeción? El argumento de Chalmers es una reiteración de la tesis que se está discutiendo.

En cualquier caso, aun si la argumentación de Chalmers fuera válida, resulta notable que en ningún momento se han dado indicaciones de en qué pueda consistir una propiedad intrínseca. Al final de la cita, Chalmers señala que las únicas propiedades intrínsecas con las que tenemos familiaridad directa son las

propiedades fenoménicas, pero esto difícilmente nos puede dar una idea de cómo podríamos llegar a saber que hemos dado con otras propiedades intrínsecas que no sean fenoménicas. El problema es que, una vez más, sigue sin estar claro si Chalmers habla de una intuición básica, reconocible desde el sentido común, o de una distinción teórica, cargada de metafísica, e introducida por la necesidad de mantener la coherencia conjunta de varios supuestos, entre ellos que es necesaria la existencia de un nivel último que sustente todo el entramado jerárquico de la realidad. Si las propiedades fenoménicas se definen, desde un principio, como incapaces de someterse a un análisis funcional, la falsedad del funcionalismo respecto a la experiencia consciente es una verdad analítica; por otra parte, según la concepción descrita, si de entrada todo lo que tenemos son propiedades funcionales, y no disponemos de una noción sustantiva de propiedad física, nos vemos abocados al problema de que no podemos comprender cómo puede haber relaciones causales en general, y esto nos lleva a la postulación de unas propiedades que no sabemos en qué podrían consistir.

La dicotomía entre propiedades disposicionales y propiedades categóricas tiene un largo recorrido en la historia de la filosofía. Las propiedades categóricas nos hablan de lo que las cosas son en sí; un ejemplo de lo que suele entenderse por propiedad categórica es la forma de un objeto. Las disposiciones, por otra parte, son propiedades problemáticas en la medida en que hacen referencia a cómo un objeto se comportaría en determinadas circunstancias; es difícil determinar las condiciones de atribución de una disposición, ya que las disposiciones parecen hacer referencia a lo que puede pasar, sin que necesariamente pase. Por ejemplo, ¿cuándo podemos decir que un objeto es frágil? Podría decirse que un vaso es frágil si, en caso de que se deje caer, se rompe; sin embargo, para que un vaso sea frágil, no es necesario que se rompa nunca, ni que sea lanzado, de modo que nada de lo que le ocurre a un vaso realmente es necesario para que se considere frágil.

Como hemos visto, comprender esta dicotomía es muy importante para la filosofía de la mente contemporánea, en la medida en que la posición funcionalista dominante distingue entre el rol causal de una propiedad y el realizador de esa propiedad, de manera que la determinación de qué poderes causales tiene un

realizador concreto es lo que, en el marco materialista esbozado, permitiría mantener la relevancia causal de lo mental. Cualquier realizador puesto fuera de un contexto relevante no puede desempeñar el rol causal adecuado que permite atribuir una propiedad funcional. Por ejemplo, un motor no puede desempeñar el rol causal adecuado si se extrae de todo el mecanismo que compone un coche. No obstante, aun si se acepta la validez del análisis funcionalista, no está claro que dicho análisis sea compatible con el carácter básico de las propiedades físicas.

Según el marco esbozado en el Capítulo 1, una explicación plausible que dé cuenta de las disposiciones es la existencia de una propiedad no disposicional que diferencie, por ejemplo, un vaso frágil de uno no frágil (para una solución de este tipo véase Armstrong 1968). Las propiedades físicas deberían ser los únicos candidatos posibles para ser la base no disposicional de las propiedades funcionales; si no, ¿en qué sentido podría decirse que las propiedades físicas son básicas? En teoría, una propiedad no disposicional, como la estructura molecular, podría explicar por qué un determinado objeto o sustancia se comportan de un modo determinado en ciertas circunstancias. Si buscar una propiedad no disposicional es la clave para entender las atribuciones de propiedades no disposicionales, el hilo argumentativo seguido por Chalmers tiene sentido y sus consideraciones respecto a la causalidad son plausibles. Sin embargo, ¿es cierto que la clave para entender qué es una propiedad disposicional es la identificación de una propiedad no disposicional que sirva de base categórica a la propiedad anterior?

David Mellor (1974) cuestionó la idea de que las disposiciones no puedan ser propiedades como las demás y analiza la idea de que las disposiciones dependan de bases categóricas. Según Mellor, una teoría que asuma que las condiciones de atribución de una disposición requieren propiedades no disposicionales tiene que poder dar cuenta de cómo las bases no disposicionales se relacionan con los contrafácticos que se asocian a la disposición. Si esta relación es contingente, entonces hay que poder individuar instancias de fragilidad con independencia de la base categórica y el criterio de atribución resulta inútil. Podría ocurrir que hubiera una conexión necesaria a posteriori entre la propiedad no disposicional y la disposición, pero en ese caso también sería necesario poder identificar la propiedad

disposicional independientemente, ya que no habría un conocimiento a priori de la conexión.

Si volvemos al ejemplo de la estructura molecular, ¿cómo dicha propiedad puede respaldar la atribución de propiedades disposicionales? En última instancia, al hablar de estructura molecular se hace referencia a ciertos contrafácticos que involucran masas y las aceleraciones de dichas masas al ser sometidas a diferentes fuerzas, de manera que la estructura molecular es una disposición entre otras; ¿pero por qué debería postularse un conjunto de propiedades intrínsecas que fundamente las disposiciones? Un argumento podría ser que las disposiciones parecen reducirse a ciertos contrafácticos, sin embargo, como acabamos de ver en el caso de la estructura molecular, no parece que la referencia a contrafácticos sea exclusiva de las disposiciones. Quizá el problema es que las disposiciones deberían poder explicar sus manifestaciones, de modo que, por ejemplo, la fragilidad debería poder explicar la rotura de un vaso al ser lanzado; pero si la atribución de fragilidad a un vaso depende de ciertos contrafácticos que hacen referencia a la rotura del vaso, parece que la explicación que puede dar la fragilidad es trivial, del mismo modo que la *vis dormitiva* explicaría por qué el opio duerme. Para evitar la trivialidad de la explicación que puede dar una disposición, es necesario que haya algún tipo de base independiente que permita su atribución; pero, según Mellor, esta base independiente puede ser otra disposición. Y lo que puede garantizar que una propiedad disposicional sea genuinamente explicativa es que se manifieste de más de una forma.⁷² Por ejemplo, si bien la estructura molecular se manifiesta en la aceleración de ciertas masas al ser aplicada una fuerza, se puede determinar una función que haga corresponder una aceleración diferente, dependiendo de las distintas fuerzas que se apliquen. Si el argumento de Mellor es correcto, no solo no hay necesidad de postular propiedades categóricas, sino que aun si contáramos con dichas propiedades, carecerían de poder explicativo.

Por otra parte, ni siquiera está claro en qué pueda consistir una propiedad categórica o intrínseca, como Chalmers señala. Al presentar el marco general, hablé

⁷² Mellor también habla del principio de conectividad como un principio de individuación de propiedades, según el cual, dos cosas no pueden diferir en una sola propiedad, de otro modo, la diferencia no podría ser explicada.

de cómo el fisicismo presupone que las propiedades físicas son básicas en algún sentido especial, pero no di ninguna definición de propiedad física por mor de la argumentación. La intuición que subyace al fisicismo es que una vez fijados los hechos físicos del mundo, el resto de hechos queda también fijado. Esta es la idea que atrapa la relación de supervinencia, en este caso, supervinencia global. En su versión local, según esta intuición, todo par de individuos que compartan sus propiedades físicas, compartirán el resto de propiedades.

El problema de esta última intuición es que, como está formulada, no nos va a permitir comprender cómo la instanciación de propiedades físicas determina o fija las propiedades funcionales, ya que, como veíamos, para que un determinado realizador mantenga las relaciones causales adecuadas que nos permitan identificar la propiedad funcional realizada, tiene que situarse en un contexto adecuado. Un motor enterrado en un desierto no puede fungir como un motor. Ahora bien, no parece que este sea un gran problema para el fisicismo; por supuesto, hay propiedades que solo pueden darse o adquirirse en determinados contextos, pero son las propiedades físicas las que fijan que un determinado objeto adquiera una determinada propiedad en cierto contexto de realización. Esta intuición se puede sintetizar en lo que Corbí y Prades (2000) llaman el Test de la Sustitutividad de los Gemelos (llamémoslo TSG): dados dos duplicados físicos, A y B, y un contexto espaciotemporal C que fije la posesión de la propiedad P por parte de A, si B estuviera en el lugar de A, B también tendría la propiedad P. De TSG podría deducirse que las propiedades físicas de un objeto fijan el resto de sus poderes causales⁷³, de manera que el drenaje de los poderes causales se habría suturado y el peligro de que el mundo acabe siendo una red de relaciones causales vacuas quedaría conjurado. Ahora bien, como veremos a continuación, es dudoso que realmente TSG pueda fijar, aunque sea indirectamente, un conjunto de propiedades intrínsecas que tenga relevancia explicativa.⁷⁴

La ventaja principal que el funcionalismo tiene frente a la teoría de la identidad de tipos es que permite la múltiple realización de propiedades mentales.

⁷³ Corbí y Prades llaman a este principio intrinsiqueza (*intrinsicness*).

⁷⁴ En el argumento que desarrollo a continuación sigo Corbí y Prades (2000).

Esto significa que, al determinar cuándo un objeto posee una propiedad funcional, hay que dejar espacio para la múltiple realizabilidad de la propiedad. Una propiedad funcional es una disposición y al decir que un objeto posee una propiedad funcional F, lo que estamos diciendo es que hay ciertos contrafácticos del objeto que son verdaderos. Estos contrafácticos podrían tener la siguiente forma:

(1) Un objeto tiene la propiedad funcional F si produjese el *output* O, si se situara en el contexto C.

Sin embargo, el tipo de contrafáctico adecuado no puede ser tan simple, ya que cada realizador de la propiedad funcional estará constituido por propiedades físicas diferentes, de manera que hacer referencia simplemente a un contexto de realización no nos permite definir adecuadamente cuándo un objeto tiene una propiedad funcional. Por ejemplo, un coche no podría ser un vehículo en la luna, puesto que no hay combustión. Para poder individuar instancias de propiedades funcionales, es necesario hacer explícitas ciertas condiciones de trasfondo que permiten que se instancie la propiedad; una bombilla no deja de ser una bombilla por el hecho de que no haya corriente eléctrica. Por tanto, el contrafáctico debería decir algo así:

(2) Un objeto tiene la propiedad funcional F si produjese el *output* O, si se situara en el contexto C y se dieran las condiciones de trasfondo CT.

La ventaja de esta definición es que puede dar cuenta de cómo en un objeto se realiza una determinada propiedad funcional en virtud de su estructura física, de manera que el requisito de la múltiple realización queda salvaguardado. (2) incluye un doble nivel de modalidad que podría considerarse prescindible, ya que CT podría incluirse en C. Sin embargo, como veremos ahora, esta maniobra no es posible. (2) no es todavía una formulación que nos permita individuar las propiedades funcionales relevantes. El problema es que el conjunto de circunstancias bajo las cuales se individua una propiedad funcional podría ser demasiado amplio, y

entonces reconoceríamos algunos objetos como poseyendo una propiedad funcional que no querríamos atribuirles; por otra parte, si fuera demasiado restrictivo, quedarían fuera objetos que querríamos reconocer como poseyendo dicha propiedad. ¿Pero cómo podemos determinar un conjunto de circunstancias relevante que no sea ni demasiado amplio, ni demasiado restrictivo para atribuir la propiedad, por ejemplo, de ser un coche? No estamos interesados en posibilidades remotas en las que un vehículo esté fabricado con hielo o con papel, lo que determina cuál es el listado de circunstancias relevantes son las circunstancias que consideramos normales:

(3) Un objeto tiene la propiedad funcional F si produciría, bajo circunstancias normales, cierto *output* O , si se situara en el contexto de realización relevante CR .

En principio, hacer referencia a unas circunstancias normales podría parecer un movimiento conveniente ajustado a nuestras prácticas reales, sin relevancia metafísica; sin embargo, parte del argumento de Corbí y Prades es que la referencia a unas condiciones normales es ineludible. Volviendo al problema del drenaje de los poderes causales, ya vimos que las dudas planteadas respecto a la eficacia causal de las propiedades funcionales podrían extenderse a casi cualquier caso de causalidad, a menos, aparentemente, que haya un nivel último de propiedades. El funcionalismo, en sí mismo, no se compromete con la existencia de este nivel último, pero según el marco fisicista esbozado, la eficacia causal de lo mental es rescatada a través de su realización física. Para que las propiedades físicas cumplan con este rol, es importante que una vez fijadas éstas respecto a un objeto, queden fijados el resto de poderes causales del mismo. Por supuesto, qué poderes causales manifiesta un objeto en un momento dado dependerá de en qué contexto se encuentre. Por mor de la argumentación, dejé de lado la cuestión de cuáles podrían ser las propiedades físicas, si bien TSG podría permitirnos fijar, en principio, un conjunto básico de propiedades. Si el análisis de las disposiciones de Mellor (1974) es correcto, no es necesario postular bases no disposicionales que garanticen la realidad y el poder explicativo de las propiedades disposicionales. Aunque no se

haya podido determinar de un modo claro un conjunto de propiedades físicas que den cuenta de los poderes causales de las distintas entidades que pueblan el mundo, al menos puede asumirse que hay un modo específico en que cada objeto, dada su realización física particular, contribuye al contexto de realización de un modo particular que le permite realizar la propiedad funcional en cuestión. Podría decirse que las propiedades físicas, independientemente de cada contexto, fijarían, para todo mundo posible⁷⁵, las posibles disposiciones de todo objeto, de manera que, al menos, debería existir una metadisposición que pudiera explicar de un modo no trivial por qué un determinado objeto tiene la disposición a ejemplificar una propiedad funcional en un contexto de realización determinado. Ahora bien, para que dicha metadisposición tenga poder explicativo⁷⁶, siguiendo a Mellor, tiene que manifestarse en más de una forma. Sin embargo, la única manera que tenemos de fijar dichas propiedades es TSG. De entrada, TSG no parece fijar una metadisposición que no sea trivial. Si el conjunto putativo de propiedades básicas está fijado exclusivamente por TSG, ¿cuál sería la manifestación independiente de dichas propiedades básicas? Dos objetos que compartieran todas sus propiedades físicas compartirían el resto de sus propiedades, y esta sería una manifestación de dicha disposición, pero ¿qué podría contar como una manifestación adicional? Al igual que en el ejemplo de la estructura molecular podía definirse una función que hiciera corresponder la aplicación de ciertas fuerzas con ciertas aceleraciones, quizá podría establecerse una función que describiera las distintas manifestaciones de las propiedades básicas de un objeto cualquiera. Esta función podría describirse como un condicional en el que, dado un subconjunto de propiedades básicas P de un objeto y un subconjunto de propiedades básicas de un contexto de realización CR y de unas condiciones de transfondo CT, respectivamente, se daría, necesariamente, un *output* O con unas propiedades básicas determinadas, de manera que:

⁷⁵ Este requisito tiene que ver con qué el fisicismo habla de cómo los hechos físicos determinan el resto de hechos, y esta determinación ha de ser válida en todo mundo posible, que es precisamente lo que los argumentos presentados niegan.

⁷⁶ Esto es lo que Corbí y Prades llaman metadisposición causal.

P & CR & CT hace necesario O

P fijaría un subconjunto de propiedades básicas que en conjunción con CR y CT garantizaría O en toda situación físicamente posible. El problema es que un condicional de este tipo no es verdadero, ya que, para cualquier P, CR y CT que se incluya en el antecedente, siempre es físicamente posible que algún factor adicional interfiera e impida que se dé O, y no importa cuán grande sea el condicional o cuántos factores se incluyan en el mismo, siempre será posible que un factor adicional impida que se genere O.⁷⁷ Por supuesto, se podría indicar que en el antecedente del condicional se incluya una cláusula que indique que ningún factor adicional impedirá que se genere O; el problema es que, en tal caso, no habríamos determinado un conjunto de propiedades básicas que fije unos poderes causales en toda situación físicamente posible. Ahora bien, si un condicional de este tipo no puede ser verdadero, entonces no se puede dar contenido a la idea de que hay unas propiedades básicas que fijan los poderes causales de las cosas.

En nuestras prácticas cotidianas, tendemos a distinguir, con bastante facilidad, entre las propiedades intrínsecas de un objeto y las propiedades que dicho objeto adquiere por estar situado en cierto contexto. Por ejemplo, algunos objetos sólidos, como los metales, mantienen la mayor parte de sus propiedades a través de los múltiples contextos en los que los seres humanos se mueven en la superficie de la Tierra, y se constituyen en haces altamente estables de propiedades. Sin embargo, nuestras intuiciones respecto a la estabilidad de los objetos metálicos sólo se dan en el contexto altamente restringido de la superficie terrestre y ni siquiera en el interior de la Tierra se mantienen estas condiciones. Las intuiciones que dan lugar al modelo jerárquico de la realidad que se ha estado discutiendo en este capítulo tienen como origen un contexto de vida bastante limitado y restringido, y cuando las intuiciones formadas en dicho marco se hipostasian como realidades metafísicas últimas dejan de ser inteligibles.

⁷⁷ El argumento no se basa en nuestros límites epistémicos. Incluso si un ser omnisciente pudiera enumerar todos los factores que de hecho intervendrían en el condicional, todavía existiría la posibilidad física de que un factor adicional falseara el condicional.

La reflexión filosófica moderna tiende a tomar la ciencia natural y, más concretamente, la física como el modelo por excelencia de una reflexión racional. Parece que la física contemporánea es la ciencia que en un menor grado involucra intereses humanos y, por ello, por las razones que vimos en el Capítulo 1, podría ser la concepción más próxima a una concepción absoluta de la realidad. Por otra parte, si la física no involucrara intereses humanos, también sería un buen punto de partida para una metafísica. No obstante, si el argumento presentado en esta sección funciona, cualquier metafísica tendría que lidiar con el problema de las condiciones de normalidad que permiten individuar determinadas propiedades. Este problema se ve mucho más claro en otras ciencias vinculadas fuertemente con la física, como puede ser la biología o la medicina. Veamos un ejemplo adicional que puede ilustrar mejor las consecuencias de la argumentación. Imaginemos una teoría farmacológica que postule que toda sustancia química activa lo es en virtud de tener una determinada propiedad intrínseca, signifique lo que signifique esto. Imaginemos una sustancia X que es mortalmente venenosa, a menos que se contrarreste con la sustancia Y, y que la explicación de estos efectos remitiera, en el caso de X, a la propiedad “ser mortalmente venenosa”, y en el caso de Y, a “ser antídoto de X”. Imaginemos, además, que dicha ciencia procediera postulando en ambas sustancias la realización de dicha propiedad en un conjunto de propiedades físicas. La propiedad “ser mortalmente venenoso” no es una propiedad intrínseca, sino disposicional. No obstante, podemos imaginar que dicha ciencia explicaría en virtud de qué otras propiedades la sustancia X es mortalmente venenosa: dicha ciencia podría apelar a unas propiedades esenciales o intrínsecas que explicaran por qué la sustancia en cuestión es venenosa en determinados casos o circunstancias. Un problema básico que tendría que afrontar la teoría es que la propiedad “mortalmente venenoso” puede ser atribuida a casi cualquier sustancia; todo depende de las circunstancias y la cantidad del consumo; incluso la ingestión de una determinada cantidad de agua en un corto período de tiempo puede producir la muerte por sobrehidratación. Por otra parte, puede decirse que en cantidades y circunstancias normales de consumo el agua no es venenosa mientras que la sustancia X sí; sin embargo, aunque legítimo, este movimiento socava cierta manera de entender las

propiedades intrínsecas, ya que, por hipótesis, la sustancia X no produce la muerte si, a su vez, se ingiere la sustancia Y, de manera que para que la teoría sea sustantiva tendría que poder realizar la distinción entre ser venenoso en virtud de una propiedad intrínseca y ser venenoso dadas ciertas circunstancias que podrían definirse como anómalas. El problema es que esta distinción no puede establecerse al margen de cierto contexto empírico y de ciertas circunstancias, de modo que nuestras atribuciones de propiedades a sustancias serían muy distintas si el organismo humano estuviera constituido de otra forma, o si lo que consideramos circunstancias normales fueran distintas (quizá la sustancia X no produciría la muerte si el nivel de oxígeno de la atmósfera fuera ligeramente superior), y la distinción no se puede llevar a cabo incondicionalmente si ha de ser inteligible. Pero si esto es así, ¿en qué medida la discusión respecto a lo que es esencial o intrínseco no se vería afectada? ¿Es posible, en última instancia, distinguir propiedades intrínsecas sin remitirnos a cómo esas propiedades se manifiestan relacionadamente?⁷⁸ Por otra parte, una química que remita a propiedades intrínsecas como términos básicos e irreductibles se encontraría muy limitada explicativamente. Por supuesto, una ciencia de este tipo podría ser muy útil; clasificar el cianuro como un veneno, en contraste con el agua o el zumo de naranja, es una información sustantiva y de gran valor práctico, pero preguntar qué es un veneno y qué no lo es, de un modo incondicional, no tiene sentido. En el caso del estudio de la mente, la tarea filosófica principal de intentar buscar un lugar para lo mental en un mundo físico que se asume que está cerrado y que es completo ha monopolizado el debate y, de este modo, se ha generado la ilusión de que es posible dar cuenta de lo mental sin tener cuenta los intereses humanos, que se manifiestan en lo que normalmente se define como condiciones de normalidad.

A continuación, analizaré una forma especial de epifenomenismo que se deriva de la concepción de lo fenoménico de Chalmers, y de cómo en su respuesta

⁷⁸ En Dupré (1993) se encuentra una crítica a las taxonomías biológicas que cuestiona la existencia de unas propiedades que, independientemente de cualquier interés investigativo resulten esenciales para determinar un género natural. Al hablar de unas propiedades intrínsecas que podamos considerar en algunos casos como esenciales, cabría preguntarse si no consideramos dichas propiedades como esenciales en virtud de cómo esas propiedades se manifiestan de un modo robusto en una multiplicidad de contextos que son relevantes para nuestros intereses científicos.

a varias objeciones se sirve de la idea de que las propiedades fenoménicas son intrínsecas. Trataré de mostrar cómo el concepto de experiencia consciente del que se sirve Chalmers, si bien puede ser coherente con el epifenomenismo y conformar una posición consistente, no es coherente con intuiciones básicas respecto a nuestras atribuciones de experiencia consciente a un ser vivo. Al igual que con el caso de las intuiciones causales analizadas en esta sección, trataré de mostrar cómo ciertas prácticas cotidianas relacionadas con nuestra forma de entender nuestras atribuciones de experiencia consciente toman en el análisis filosófico un cariz metafísico que excede la lógica de su aplicación.

4.2. La paradoja del juicio fenoménico

La cuestión del epifenomenismo planteada por la concepción de la experiencia consciente surge como un problema adicional a las dificultades discutidas. De hecho, uno de los puntos de partida de la filosofía de Chalmers es que la dimensión fenoménica de la experiencia consciente es misteriosa en un sentido en el que no lo es la dimensión psicológica, la cual coincide con el programa de la ciencia cognitiva y con el éxito, al menos en principio, de la concepción funcionalista de lo mental. El tipo de proyecto que trata de emprender Chalmers se caracteriza, desde sus inicios, por una tensión cuya resolución es difícil, en la medida en que aparece en los mismos fundamentos de su planteamiento. Por una parte, cualquier criterio público de experiencia consciente fenoménica es conceptualmente independiente de ésta última y, en última instancia, el que la fenomenología constituya un *explanandum* depende del conocimiento íntimo y privado que cada individuo tiene de ésta. Por otra parte, la misma reflexión sobre la experiencia consciente tiene una pretensión de objetividad y de comunicabilidad, muy especialmente en el caso de Chalmers, quien pretende esbozar unas bases de lo que debería ser una ciencia de la conciencia. Es habitual encontrar al inicio de una reflexión sobre la conciencia toda una serie de casos en los que indefectiblemente atribuimos experiencia consciente y respecto a los que, en un sentido práctico, no gozamos de serias dudas. Por ejemplo, al reflexionar sobre la experiencia consciente de un murciélago

son muchos los interrogantes que se plantean en relación a las grandes diferencias que pueda haber entre la experiencia consciente de un ser humano y la de éstos, dadas las diferencias fisiológicas entre ambos, pero parece que el hecho de que los murciélagos gocen de experiencia consciente está más allá de toda duda. No obstante, en la historia de la filosofía no hay unanimidad respecto a estas intuiciones. Descartes consideraba que los animales eran autómatas que, en última instancia, no se diferenciaban en lo relevante de los mecanismos complejos. En cualquier caso, las razones metafísicas y teológicas que apoyaban esta hipótesis no se encuentran vigentes en el debate actual y, por otra parte, el desarrollo de la biología moderna, tanto en lo que concierne a la teoría de la evolución como a la comprensión de la biología molecular, ha llegado a hacer bastante implausible la idea de que haya diferencias ontológicas y metafísicas entre el ser humano y otros seres vivos.

En lo que resta de capítulo, mostraré el profundo abismo generado entre las prácticas de la vida real concernientes a las atribuciones de conciencia y los debates metafísicos en torno a la misma, y en qué medida la confusión entre ambas esferas distorsiona el debate. En esta sección del capítulo, centraré mi atención en el contenido de la paradoja del juicio fenoménico, y dedicaré la sección 4.3 al análisis de las atribuciones prácticas de experiencia, tratando de mostrar que la paradoja es ininteligible, aun si no se tienen en cuenta algunas objeciones generales vinculadas a la cuestión de la causalidad.

4.2.1. Los tipos de juicios fenoménicos

Los juicios fenoménicos tratan de la dimensión fenoménica de lo mental. No obstante, al hablar de juicios, hay que tomar cierta decisión previa respecto a qué se considera como un juicio o como una creencia; en este sentido, Chalmers toma como punto de partida la concepción de juicio como un estado funcional, individuado por sus vínculos causales con el comportamiento, el entorno y otras creencias. En segundo lugar, Chalmers deja abierta, inicialmente, la cuestión de si las propiedades fenoménicas determinan el contenido de las creencias y los juicios fenoménicos.

Por otra parte, es importante recordar que según Chalmers, por distintos que sean conceptualmente, los criterios de los que disponemos para atribuir estados conscientes a otros seres se solapan con los criterios vinculados a estados cognitivos. El objetivo de Chalmers es sentar las bases de una teoría de la conciencia que tenga una aplicación científica y, para ello, los nexos entre cognición y fenomenología son fundamentales:

Alguien podría tener la impresión de que una teoría de la conciencia y una teoría de la cognición tendrán poco que ver la una con la otra.

Esta imagen es engañosa. Nuestra vida mental no está alienada de sí misma de la manera que la imagen sugiere. Hay vínculos profundos y fundamentales entre conciencia y cognición. Por una parte, los contenidos de nuestras experiencias conscientes están estrechamente relacionados con los contenidos de nuestros estados cognitivos. Siempre que alguien tiene una sensación de verde, individuada fenoménicamente, tiene una correspondiente *percepción* de verde, individuada psicológicamente. Por otra parte, gran parte de la actividad cognitiva puede centrarse en la experiencia consciente. Sabemos de nuestras experiencias y hacemos juicios sobre ellas; mientras escribo esto, gran parte de mi pensamiento está siendo dedicado a la conciencia. Estas relaciones entre conciencia y cognición no son arbitrarias ni caprichosas, sino sistemáticas. (Chalmers 1996: 172)

Chalmers distingue tres tipos de juicios fenoménicos para dar cuenta de la ambigüedad. Los juicios fenoménicos de primer orden se caracterizan porque acompañan a una experiencia consciente, pero no se refieren a la experiencia como tal, sino al objeto de la experiencia. Se trata de estados cognitivos, por lo que, estrictamente, forman parte del objeto de la psicología. Un ejemplo de estos juicios es “Este libro es rojo”, e incluyen, en general, las percepciones. Los juicios de segundo orden, que son los que más interesan a Chalmers, tienen como objeto experiencias conscientes propiamente dichas, de manera que las sensaciones son el ejemplo paradigmático, como en el juicio “Estoy teniendo una sensación de rojo”, y pueden incluir también tipos de experiencias, como cuando alguien habla de lo agradable que es el cosquilleo previo a un estornudo. Finalmente, los juicios de

tercer orden tienen en cuenta a las experiencias conscientes como un tipo, e incluyen juicios como “Las sensaciones son misteriosas” o muchos de los juicios que componen *The Conscious Mind*.

Respecto a los dos primeros tipos de juicios, recordemos que, en el Capítulo 2, se cuestionaba la posibilidad de individuar independientemente contenidos fenoménicos y psicológicos; dejaré de lado, de momento, dicho argumento. En cualquier caso, el solapamiento de los contenidos cognitivos con los contenidos fenoménicos genera una ambigüedad notable que, en algunos casos, tiende a convertirse en un error categórico. Si atendemos a la idea de Chalmers según la cual se tiene, simultáneamente, tanto una percepción de verde como una sensación verde, habría que señalar que resulta un poco extraño decir que una sensación es verde, al menos en el mismo sentido en que lo es un objeto corriente. Los colores son, normalmente, propiedades de objetos con cierto tamaño, superficie y con capacidad de reflejar la luz, pero no es éste el caso de una sensación⁷⁹. Pensemos en un ejemplo paradigmático de sensación como el dolor, para el que no existe una distinción paralela. Aun si aceptamos la distinción, habría que decir algo respecto a qué significa, en cada caso, el predicado usado. De todos modos, es una distinción que no puede aceptarse como mero sentido común, sino que está cargada ya metafísicamente.⁸⁰

Por otra parte, si atendemos a los juicios de tercer orden, no sólo se heredan algunos de los problemas referentes a los juicios de segundo orden (pensemos en el juicio “Algunas experiencias conscientes son verdes”) sino que no queda muy claro cuál es la extensión de las experiencias conscientes como tipo, para cuya delimitación sólo existe un catálogo de elementos heterogéneos, en el mejor de los casos. Este problema será especialmente apremiante si los juicios de tercer orden tienen que convertirse en la fuente de una ciencia de la experiencia consciente. Sin

⁷⁹ Hay algunas excepciones interesantes respecto a los colores. El color de las luces, por ejemplo, no puede atribuirse a los objetos que ilumina.

⁸⁰ Para una discusión detallada de los problemas de esta distinción es interesante el análisis llevado a cabo en Stroud (2000) de la distinción entre “ver proposicional” (propositional seeing) y “ver predicativo” (predicative seeing).

embargo, es a continuación cuando presentaremos el problema central que plantean los juicios fenoménicos.

4.2.2. La paradoja

La paradoja de los juicios fenoménicos consiste en que los juicios fenoménicos de segundo y tercer orden tienen por objeto la experiencia fenoménica, la cual es irreducible si los argumentos de Chalmers son válidos, pero son, a su vez, estados cognitivos y reducibles, por tanto, a estados funcionales, de manera que la experiencia fenoménica no juega ningún rol explicativo en su emisión. Intuitivamente, podría decirse que la experiencia fenoménica tampoco goza de ninguna relación causal con los juicios fenoménicos, aunque las cautelas de Chalmers respecto al debate de la causalidad, y algunas de sus consideraciones más especulativas respecto al papel que podría jugar la experiencia fenoménica en el conjunto de la realidad le llevan a ceñir la discusión al ámbito de la explicación.

La paradoja se podría formular de la siguiente manera:

- (1) Los juicios fenoménicos son sobre la experiencia consciente.
- (2) Los juicios son reducibles funcionalmente.
- (3) La experiencia consciente no es reducible funcionalmente.
- (4) Por lo tanto, la experiencia consciente es explicativamente irrelevante respecto a los juicios fenoménicos.

La manera más intuitiva de plantear la paradoja es a través del experimento mental de los zombis. Imaginemos al zombi gemelo de Chalmers, a quien podemos llamar Chalmerz.⁸¹ Chalmerz afirmaría que goza de experiencias conscientes y

⁸¹ Se nos ocurren dos posibilidades para plantear el experimento: podemos intentar imaginar que Chalmerz habita una Tierra Gemela zombi, o en una Tierra Gemela en la que la única variación es que Chalmerz es un zombi.

podría dar descripciones detalladas de las cualidades de las mismas, indistinguibles, en sus rasgos generales, de las descripciones que daría cualquier persona. Además, Chalmerz habría reflejado todas sus ideas en un libro llamado *The Conscious Mind*, y habría argumentado a favor de su irreductibilidad; ahora bien, todos estos hechos podrían explicarse, exclusiva y exhaustivamente, desde el punto de vista de una ciencia cognitiva completa. Pero lo más importante es que si todos los hechos mentales relativos a Chalmerz pueden explicarse cognitivamente, no hay ninguna diferencia a ese respecto entre Chalmers y Chalmerz, ya que Chalmers es un reflejo isomórfico de Chalmerz. Igualmente, todos y cada uno de los movimientos que llevamos a cabo podrían explicarse sin referencia a la experiencia consciente.

Como ya adelanté, la paradoja se genera a partir de la tensión entre la clausura causal del mundo físico y la irreductibilidad de la experiencia consciente, de manera que se plantea un dilema semejante al que apunta Kim cuando describe la disyuntiva: “o bien reducción, o bien impotencia causal” (Kim 2006: 54). Por otra parte, según Chalmers, no se puede renunciar plausiblemente a la clausura causal del mundo físico, ni a la irreductibilidad de la experiencia consciente y, aceptando que los argumentos presentados funcionen, tampoco puede negarse el carácter de *explanandum* de la experiencia consciente, de manera que, a lo sumo podemos aprender a vivir con la paradoja. Pero esta consecuencia no es tan preocupante ya que, según Chalmers, la aceptación de los principios que generan la paradoja no conduce a una posición incoherente, por muy contraintuitiva que resulte. Sin embargo, Chalmers tiene en cuenta algunos argumentos que sí parecen poner en cuestión el punto de vista asumido y, para responder a estas cuestiones, profundizará más en el tipo de relación que mantenemos con nuestros propios estados mentales.

4.2.3. Experiencia consciente, relevancia causal y familiaridad

La paradoja del juicio fenoménico no se limita a ser problemática desde el punto de vista explicativo. Parece un hecho incontrovertible que conocemos, recordamos y

nos referimos a nuestras propias experiencias conscientes. Sin embargo, ¿cómo puede explicarse el conocimiento que tenemos de nuestra propia fenomenología? Si tanto los juicios fenoménicos proferidos por mí como los juicios paralelos de mi zombi gemelo pueden ser explicados por apelación a los mismos procesos cognitivos, entonces los juicios y creencias de mi zombi gemelo estarían justificados en la misma medida en la que lo están los míos, pero en tal caso mis juicios no estarían justificados; de hecho, los juicios de mi zombi gemelo son sistemáticamente falsos.⁸² Por otra parte, parece razonable pensar que, para que haya conocimiento, es necesario que se den ciertos vínculos causales con lo que se conoce. Pero si se acepta la irrelevancia causal de la conciencia, la única alternativa que le resta a Chalmers para no rechazar el conocimiento de nuestra fenomenología es negar que sea necesario un vínculo causal para que haya conocimiento. Una posibilidad de seguir esta vía sería explicar el conocimiento de la experiencia consciente desde una posición fiabilista. Las creencias de un zombi no pueden estar justificadas en la medida en que lo están las de un ser humano normal, porque los juicios y creencias de un zombi son sistemáticamente falsos, y el mecanismo que les lleva a emitir sus juicios y formar creencias no es un proceso fiable que dé lugar a creencias y juicios verdaderos. Sin embargo, según Chalmers, el tipo de conocimiento que tendríamos de nuestra propia experiencia consciente no sería, en tal caso, el adecuado: el problema es que, así como una concepción fiabilista del conocimiento puede dar cuenta, en principio, de nuestro conocimiento perceptivo de objetos físicos, de acuerdo con Chalmers, no se puede no tener certeza respecto a si uno goza de experiencia consciente o, al menos, no en un sentido paralelo respecto al de la duda en la existencia de objetos físicos. La idea es que, en relación a la experiencia consciente, no hay espacio conceptual para introducir escenarios escépticos: en principio, tendría sentido la hipótesis de que los objetos que veo no son reales, sino el mero resultado de unos *inputs* suministrados a un cerebro en una cubeta, ¿pero en qué consistiría una posibilidad paralela en la que se engaña a un sujeto en cuanto

⁸²Chalmers presupone una noción deflacionaria de creencia y de juicio que se reduce al desempeño de un rol causal, y no presupone la fenomenología como un componente necesario para determinar el contenido de éstos. Por otra parte, en relación a la idea de justificación, cabe señalar que un juicio podría ser falso y estar justificado, pero no es posible que esté justificado si es sistemáticamente falso.

a que tiene experiencia consciente aun cuando éste no la tiene de hecho?⁸³ Esto no sería posible conceptualmente, en la medida en que “la experiencia consciente se encuentra en el centro de nuestro universo epistémico. Tenemos acceso a ella directamente”. Dicho de otro modo, “la experiencia es parte de nuestra situación epistémica central”,⁸⁴ de manera que un escenario en el que varíe nuestra experiencia consciente ya no puede ser un escenario en el que el sujeto tenga la misma evidencia. Por otra parte, Chalmers prosigue, el conocimiento de la propia experiencia consciente no puede depender de un vínculo contingente como el de la relación causal, puesto que, al igual que en el caso de la relación fiabilista, habría un hueco conceptual en el que cabrían escenarios escépticos. En este sentido, negar el vínculo causal que fundamentaría el conocimiento de nuestra experiencia consciente no sólo es una alternativa para Chalmers, sino que es una consecuencia necesaria de su posición. Ahora bien, en tal caso, ¿qué tipo de vínculo podría dar cuenta de la relación con nuestra propia experiencia consciente? Esta es la respuesta de Chalmers:

Todo esto es como decir que hay algo intrínsecamente epistémico respecto a la experiencia. Tener una experiencia es, de inmediato, encontrarse en algún tipo de relación epistémica íntima con la experiencia—una relación que podríamos llamar “familiaridad” (*acquaintance*). No hay, incluso, una posibilidad conceptual de que un sujeto pudiera tener una experiencia de rojo como ésta sin tener *algún* contacto epistémico con ella: tener la experiencia es relacionarse con ella de esta manera. Nótese que no digo que tener una experiencia sea automáticamente *saber* de ella, en el sentido en que el conocimiento requiere creencia. Creo que esa tesis sería falsa: tenemos muchas experiencias respecto a las que no tenemos creencias, de manera que no sabemos acerca de ellas. Además, se podría tener una experiencia sin conceptualizar dicha experiencia de ningún modo. Tener una experiencia y, consecuentemente, estar familiarizado con la experiencia consiste en mantener una

⁸³En última instancia, estas son las mismas intuiciones de Descartes cuando concluye que hay un límite a partir del cual ya no se puede dudar, a saber, cuando se trata de los propios pensamientos, término que en su contexto coincide, parcialmente, con los aspectos fenoménicos de la experiencia consciente.

⁸⁴ Chalmers (1996: 182).

relación con ésta más primitiva que la creencia: proporciona *evidencia* para nuestras creencias, pero no constituye en sí mismo creencia.

De hecho, nada de lo que he dicho implica que todas las creencias sobre experiencias sean incorregibles, en el sentido en que tal creencia está, automática y completamente justificada. Puesto que las creencias sobre experiencias se encuentran a cierta distancia de las experiencias, pueden formarse por todo tipo de razones y a veces se formarán creencias injustificadas. (Chalmers 1996: 196-197)

Ya vimos más arriba que, debido a sus compromisos funcionalistas, Chalmers se veía obligado, al menos de un modo negativo, a reconocer la existencia de un tipo de propiedades que no fueran relacionales en el sentido en que lo son las propiedades funcionales. En un sentido no determinado, las propiedades fenoménicas de la experiencia gozaban de este carácter intrínseco, y esto se vinculaba directamente con la idea de que la fenomenología no puede determinarse en virtud de desempeñar un rol causal. Una vez asumidos estos supuestos, con independencia de que se tenga en cuenta la paradoja del juicio fenoménico, es inevitable que se sigan consecuencias respecto al modo de concebir la epistemología de la experiencia consciente. Sin embargo, tratar de seguir coherentemente un hilo dialéctico no implica que se estén dando justificaciones de las distintas tesis que se van aceptando. De hecho, Chalmers no establece de un modo claro en qué puede consistir la familiaridad, de manera que surgen una multitud de problemas conceptuales.

En el texto citado, Chalmers se desmarca de la idea de que la familiaridad se convierta, sin más, en conocimiento y, además, negará que gocemos de algún tipo de incorregibilidad respecto a nuestras creencias fenoménicas, ya que, por ejemplo, podemos estar distraídos respecto a nuestras propias experiencias. Esto introduce un compromiso fuertemente realista, no sólo respecto a los estados mentales de los demás, sino respecto a nuestra propia experiencia consciente. Pero no está muy claro cuál puede ser el contenido concreto de estas tesis epistémicas, en la medida en que presuponen una concepción ambigua de creencia y de juicio. Chalmers ya había advertido de que se serviría de una noción deflacionaria de estos últimos que

se vincularía a su carácter de estados cognitivos⁸⁵, de manera que quedaba abierta la cuestión de si la fenomenología influía en el contenido de nuestros estados intencionales. Sin embargo, si los estados cognitivos que constituyen las creencias y juicios fenoménicos no pueden ser conocimiento, resulta incomprensible entender cómo el añadido de la experiencia consciente puede convertirlos en conocimiento o, en qué medida, puede llegar a ser relevante la diferencia entre estados intencionales deflacionarios e inflacionarios.

El problema no es que se haya perdido un vínculo causal necesario entre objeto y conocimiento, sino que no se ha dado una caracterización de en qué pueda consistir un vínculo alternativo que justifique y convierta en conocimiento simples estados cognitivos. En este sentido, de nada sirve reiterar que la justificación que proporciona la experiencia consciente es de otro tipo, o que es directa, si el único contenido del que están dotadas estas tesis se deriva de la mera caracterización negativa de las relaciones que no puede mantener la experiencia consciente con los estados cognitivos. Una teoría que afirmara que gozamos de un órgano no material que nos permite percibir estados internos de un modo introspectivo e infalible sería una tesis ciertamente falsa. Sin embargo, la concepción de Chalmers tampoco puede dar cuenta de una manera clara del error respecto a la experiencia consciente. ¿Qué puede querer decir que respecto a nuestra fenomenología tenemos una relación más básica que la de creencia, pero que las creencias que más tarde formemos sobre nuestra experiencia son falsas? ¿Qué sentido puede tener la idea de que la experiencia fenoménica constriñe o distingue nuestras creencias de un modo relevante para considerar dicha relación como conocimiento o justificación?⁸⁶

⁸⁵ En este sentido, los zombis gozan de creencias y emiten juicios.

⁸⁶ Obsérvese que este problema es distinto de lo que Wilfrid Sellars bautizó como el mito de lo dado. El problema planteado por lo dado, que se discute de un modo profundo en McDowell (1994), es cómo la experiencia entendida como carente de articulación conceptual puede servir de justificación racional de nuestro conocimiento y de nuestras creencias, y no de una mera causa bruta de éstos, y de cómo, en general, una experiencia carente de articulación conceptual puede tener una influencia inteligible en estados dotados de un contenido conceptual. Ésta es una cuestión que Chalmers no trata de responder, y en la nota sexta del capítulo quinto de *The conscious Mind* se limita a señalar que no es un problema que de un modo especial surja respecto al no reduccionismo de la experiencia consciente. El problema que ahora se discute no puede plantear la dicotomía entre que la experiencia sea una causa bruta de nuestras creencias o que goce ya de un

Hay otros argumentos que se construyen en torno a la idea de que son necesarios vínculos causales para que podamos recordar nuestra experiencia consciente, así como para referirnos a ésta, y algunas de las cosas dichas se pueden aplicar paralelamente. En el caso de la memoria, hay un sentido psicológico según el cual cuando alguien emite juicios respecto a sus experiencias pasadas, hay una explicación cognitiva de dichos juicios; sin embargo, sólo podrá hablarse de memoria en un sentido específico, frente a la pseudomemoria de los zombis, cuando se dan las experiencias conscientes oportunas y el sujeto mantiene una relación adecuada de familiaridad (*acquaintance*) con éstas. Respecto a la posibilidad de la referencia a nuestra fenomenología, pueden aplicarse consideraciones semejantes, de manera que se cuestiona la idea de la necesidad de vínculos causales para que pueda haber referencia. La discusión se complica en la medida en que Chalmers tiene en cuenta el modo en que las intensiones primarias, que son las relevantes en el caso de la conciencia, pueden referir sin la mediación de vínculos causales. En cualquier caso, la relevancia de las denominadas intensiones primarias, como vimos en el Capítulo 3, depende de la verdad del principio metafísico según el cual no es posible, en el caso de la conciencia, establecer una diferencia entre el ser y el parecer. Dicho en otro contexto y enunciado de otra forma, este movimiento dialéctico es equivalente a la afirmación de que respecto a la experiencia consciente no hay posibilidad de plantear escenarios escépticos, que en última instancia es el contenido negativo al que se reduce la idea de familiaridad.

En esta sección, he tratado de mostrar que el modo de relacionarnos con nuestra propia experiencia consciente propuesto por Chalmers para dar cuenta de la aparente incoherencia planteada por la paradoja del juicio fenoménico no se caracteriza de un modo positivo y la convierte en una maniobra *ad hoc* exigida por la dirección de la discusión. Sin embargo, no puede decirse que se haya propuesto un argumento que realmente bloquee la posición de Chalmers. En la última sección de este capítulo esbozaré una línea de argumentación que tratará de mostrar el

contenido conceptual, porque se ha renunciado a que la experiencia consciente pueda tener un vínculo causal bruto con nuestras creencias, pero también se niega que ésta tenga, de entrada, una articulación conceptual.

conflicto de las posiciones de Chalmers con las condiciones de atribución de experiencia consciente que en los restantes capítulos desarrollaré con mayor detalle.

4.3. Epifenomenismo y atribuciones de experiencia consciente

Hasta ahora, en lo que se lleva dicho respecto al epifenomenismo, no se ha tenido en cuenta el funcionamiento real de las prácticas y hechos vinculados a nuestras atribuciones de experiencia consciente. La teoría de Chalmers, a pesar de presentarse como una especie de naturalismo que pretende sentar las bases de una ciencia de la conciencia, no recurre demasiado a los datos empíricos ni a las investigaciones científicas que estudian la conciencia. Esta actitud es coherente con la idea de que, hasta ahora, la ciencia no ha dado un solo paso en la dirección de encontrar una solución para el problema de la conciencia. Sin embargo, es reveladora la escasa información que es necesaria para generar el problema de la conciencia, y la rapidez con la que se comprende.⁸⁷ Por otro lado, muchas de las tesis defendidas por Chalmers pueden verse afectadas por los resultados de la información empírica, e incluso, como trataremos de mostrar, pueden verse desmentidas por nuestras prácticas reales. Por ejemplo, si fuera cierta la tesis de que aprendemos que los demás gozan de experiencia consciente a través de una asociación entre nuestros estados mentales y nuestros movimientos corporales, de manera que en los demás inferimos la existencia de estados mentales semejantes a los nuestros, en el desarrollo infantil debería poder apreciarse una forma de aprendizaje que refleje esta forma de aprender que los demás tienen una mente o que sea, al menos, compatible con este modelo, lo cual está lejos de establecerse.

Uno de los principios que hasta ahora se han estado usando de una manera implícita es que tiene sentido hablar de fragmentos separados de experiencia consciente, así como de fragmentos de conducta intencional con cierta unidad

⁸⁷ Resulta notable que ya en el siglo XVIII Leibniz pudiera plantear el problema de la conciencia de una manera aproximada a como se plantea actualmente, aunque con un cariz mecanicista. Por otra parte, resulta todavía más interesante que el problema de la conciencia no se planteara en una forma semejante a la que conocemos, mientras no se hubo desarrollado suficientemente la filosofía moderna.

independiente. En última instancia, si el funcionalismo es verdadero, tiene sentido la idea de que podemos individuar procesos causales, o conjuntos de procesos causales, que en sí mismos son idénticos a un tipo de estado mental. Por supuesto, si aceptamos la argumentación de Chalmers, los fragmentos de experiencia consciente se corresponderían o correlacionarían de un modo natural con estados funcionales, de manera que de un modo indirecto podríamos individuar estados mentales fenoménicos a través de la identificación de la propiedad funcional con la que se corresponden. Sin embargo, parece que esta idea, sin matices, es muy problemática. Por ejemplo, suponiendo que dispusiéramos de una definición funcional correcta de dolor, ¿podríamos desarrollar un artefacto que ejemplificara una vida mental reducida exclusivamente a la implementación de estados dolorosos? Tomando un ejemplo manido en la bibliografía sobre el tema, si el dolor fuera idéntico a un determinado tipo de estimulación de las fibras C, ¿podría un conjunto separado de fibras C, estimulado de la manera adecuada, ejemplificar un estado de dolor? Una objeción que se impone a estas preguntas escépticas es que caricaturizan la concepción funcionalista de la mente; al fin y al cabo, el funcionalismo no se compromete con una idea determinada respecto a lo complejos que tengan que ser los conjuntos de relaciones causales que constituyen un estado mental, e incluso podría ser esencial a la existencia de todo estado mental la interacción con otros estados mentales. Sin embargo, dado que esta complejidad de lo mental parece esencial, los modelos de definición funcional tendrían que cambiar mucho, y no tendría mucho sentido hablar por separado de la definición funcional de dolor. Por otra parte, si la ejemplificación de las propiedades funcionales requiriera de un conjunto de relaciones causales excesivamente grande, el funcionalismo podría llegar a ser, desde un punto de vista práctico y explicativo, una teoría poco útil en el mejor de los casos y, por su falta de concreción, trivialmente verdadera.

Si tenemos en cuenta el conjunto de prácticas vinculadas a la atribución de experiencia consciente, podemos observar una gran variedad de actitudes y reacciones ante la evidencia que nos proporcionan, no sólo el comportamiento de otras personas, sino el comportamiento animal, o el funcionamiento de mecanismos

inanimados. En este sentido, si nos fijamos arbitrariamente en determinados segmentos de la conducta y tratamos de establecer correlaciones respecto a la reacción que producen en otros individuos, podemos llegar a encontrar una ausencia total de regularidad. Por ejemplo, el fragmento de conducta constituido, en cada caso, por un reflejo, un gesto de dolor, un gesto ofensivo, un espasmo o el movimiento de un brazo mecánico pueden llegar a ser, desde un punto de vista macroscópico, indistinguibles para un observador normal, a pesar de que la reacción inmediata que provocarán, cada uno de ellos, será muy distinta. Del mismo modo, entendemos perfectamente la diferencia entre la conducta de una persona despierta y la conducta de un sonámbulo, a la cual no atribuimos experiencia consciente, aun pudiendo ser esta última muy compleja.

En el ensayo de Thomas H. Huxley *On the Hypothesis That Animals Are Automata, and Its History* (1874) se describen algunos experimentos extremadamente crueles e inaceptables actualmente, llevados a cabo en ranas, en los que, a partir de ciertas lesiones infringidas en su sistema nervioso, se estudian los patrones resultantes de su conducta. Por ejemplo, se describe el caso de una rana a la que se le ha seccionado su médula espinal en la mitad de su cuerpo, de modo que resulta imposible, en principio, que sienta dolor en la mitad inferior de su cuerpo; no obstante, cuando le es aplicado ácido en la parte insensible de su cuerpo, lleva a cabo los movimientos propios de sufrir un gran dolor siendo indistinguible en ese caso de una rana sana, y trata de frotarse en la zona donde se le ha aplicado el ácido con sus patas traseras, supuestamente paralizadas. ¿Puede decirse que en tal caso la rana siente dolor?

En el caso de que una rana nos parezca un caso de una vida animal demasiado lejana a la nuestra para que podamos decir algo respecto a las intuiciones que suscitan estos experimentos, en el ensayo de Huxley se narra el caso de un soldado que sufre una importante lesión cerebral en una contienda, y que como consecuencia de la lesión sufre de ataques periódicos en los que, a pesar de mostrar una clara disfunción cognitiva, que parece incompatible con un estado de consciencia, sigue llevando a cabo una gran cantidad de acciones y movimientos que habitualmente consideramos actividad consciente. Por ejemplo, situado en un

lugar en el que se encuentra habitualmente se mueve con normalidad, aunque si se encuentra con obstáculos imprevistos choca con ellos y los evita después de tantearlos manualmente; es capaz de beber, comer, vestirse y desvestirse, aunque parece completamente insensible al dolor y es capaz de beber un vaso de vinagre como si fuera agua; el único sentido que parece conectarlo con el mundo, por otra parte, es el tacto. En este caso, ¿diríamos que dicho hombre goza de experiencia consciente? Y, en caso afirmativo, ¿consideraríamos que sólo goza de experiencias conscientes de tacto mientras que el resto de sus movimientos son automáticos, o bien diríamos que goza de una experiencia consciente plena, aunque la mayor parte de sus movimientos se sustraen a su voluntad?

Si traemos a colación estos ejemplos, no es para mostrar que podemos determinar en todo caso con certeza la existencia o ausencia de experiencia consciente, algo que claramente no podemos hacer.⁸⁸ Lo que estos ejemplos muestran es que, por una parte, no hay patrones de conducta que asociemos necesariamente a la existencia de experiencia consciente y, por otra parte, nuestras intuiciones respecto a los estados conscientes de otros seres se ven afectadas por los cambios generales que estos puedan sufrir en su comportamiento y en el conjunto de sus movimientos y reacciones corporales. Ahora bien, en ese caso, ¿cómo se pueden explicar estas intuiciones desde un modelo que presupone la correlación entre tipos delimitados de estados cognitivos y estados conscientes en el sentido fenoménico? Del mismo modo, una serie de complejos patrones de comportamiento podría hacernos creer que nos encontramos ante un ser plenamente consciente, y revelarse en un pequeño detalle como un burdo artefacto programado para llevar a cabo ciertas rutinas automáticas. Teniendo en cuenta estas consideraciones, parece que la cuestión del epifenomenismo no puede ser adecuadamente entendida según un modelo que delimita de un modo segmentado qué cuenta como una instancia de un determinado tipo mental, en la medida en que distorsiona, de entrada, lo que de un modo intuitivo y práctico entendemos por poseer atributos mentales. Entre otras cosas, poseer una mente consciente implica

⁸⁸ Un ejemplo extremo de nuestra falibilidad en estas atribuciones son los casos de personas que al entrar en coma se las ha declarado como inconscientes, y tiempo después se ha podido mostrar que no era así.

mostrar cierta coherencia y estabilidad en las expresiones, acciones, comportamientos y manifestaciones corporales, y las diferencias entre lo mental y lo no mental, y lo consciente y lo no consciente, son necesariamente difusas. Si esta idea es correcta, la experiencia consciente se encuentra ligada de una manera inextricable a nuestras acciones, y no tiene sentido llevar a cabo la separación metafísica que la posibilidad del epifenomenismo exige. En el siguiente capítulo, trataré de dar un contenido más concreto a la idea de que no tiene sentido dotar de un carácter metafísico último a la idea de que hay fragmentos aislados de experiencia consciente.

5. Experiencia consciente, realismo y el Teatro Cartesiano

En el Capítulo 3, argumenté que la concepción de la experiencia fenoménica empleada por Chalmers se caracteriza por estar cargada metafísicamente, de manera que, aun si ésta se encuentra sustentada en intuiciones básicas o en nociones del sentido común, la argumentación de Chalmers necesita principios adicionales para generar sus conclusiones. Algo que, aparentemente, convierte a la experiencia fenoménica en una dimensión única no sólo de lo mental, sino del mundo, es que respecto a las propiedades fenoménicas de la experiencia no hay diferencia entre su ser y su parecer, de manera que, por ejemplo, no es posible sentir un dolor sin que dicho dolor exista. En el Capítulo 4, presenté la relación de familiaridad (*acquaintance*) como el tipo de relación que recoge adecuadamente el fundamento del conocimiento de nuestra propia experiencia consciente; esta familiaridad se puede definir, de un modo mínimo, como vinculada a la idea de que no es posible plantear escenarios escépticos respecto a la experiencia consciente. Superficialmente, esta relación de familiaridad podría parecer un modo de dar un contenido más concreto a la coincidencia entre ser y parecer de lo fenoménico.

No obstante, la aceptación de estos principios plantea algunos problemas. Por una parte, si no es posible formular argumentos escépticos respecto a la experiencia fenoménica, parece que el conocimiento de nuestra propia experiencia fenoménica es infalible; esta concepción de la experiencia, lejos de ser iluminadora, necesita ser explicada en la medida en que plantea una asimetría respecto al resto de hechos que conocemos que no ha sido caracterizada de un modo sustantivo, sino más bien de un modo *ad hoc*. Por otra parte, en la bibliografía sobre la experiencia consciente existen argumentos que tratan de mostrar que algunas de las intuiciones más robustas respecto al carácter de la experiencia consciente son el resultado de una ilusión, atacando directamente el tipo de principio descrito más arriba. Éste es el caso de las llamadas posiciones eliminacionistas, entre las que la obra de Daniel Dennett⁸⁹ constituye uno de los intentos más articulados y completos de desarrollar este tipo de posición.

⁸⁹ Ver Dennett (1991), para otras posiciones eliminacionistas ver Churchland (1981) y Rey (1983).

En este capítulo, argumentaré en contra de la concepción metafísica que Chalmers tiene de la conciencia. En primer lugar, analizaré la posición que la argumentación de Daniel Dennett ocupa en dicho debate y me centraré en la cuestión de cuál es el tipo de conclusión que trata de alcanzar, llevando a cabo una interpretación particular de su crítica a lo que él llama la concepción del Teatro Cartesiano; a partir de esta interpretación, presentaré una argumentación en contra de la concepción de la experiencia consciente de Chalmers. Seguidamente, señalaré los límites que tiene el tipo de argumentación presentado en este capítulo.

5.1. Experiencia consciente, apariencia y realidad

He tratado de mostrar que la concepción de la experiencia consciente de la que se sirve Chalmers, lejos de ser una intuición básica, constituye una concepción metafísica que no se encuentra claramente sustentada en argumentos adicionales. De hecho, el aspecto fenoménico de la experiencia no se puede caracterizar de un modo que no plantee ya el problema de la experiencia consciente. Chalmers apunta a algunas intuiciones básicas que dan lugar, necesariamente, al problema de la conciencia. Sin embargo, esta clase de circularidad no supone en sí una objeción demasiado contundente para la argumentación de Chalmers, quien incluso es consciente de la misma:

Podría objetarse, adicionalmente, que los argumentos que he dado consisten, al final, en una colección de intuiciones. Ciertamente, hay un sentido en el que todos estos argumentos se basan en la intuición, pero he intentado aclarar simplemente cuan naturales y sencillas son estas intuiciones, y cuan forzado es negarlas. La intuición principal en juego es que *hay algo que tiene que ser explicado*—algún fenómeno asociado con la experiencia en primera persona que introduce un problema que no se presenta por la observación de la cognición desde el punto de vista de la tercera persona. [...]

La “intuición” que aquí está en juego es la misma *raison d’être* del problema de la conciencia. La única manera consistente de evitar las intuiciones es negar el problema y el fenómeno completamente. Siempre se puede, al menos cuando se

habla “filosóficamente”, negar completamente las intuiciones y negar que haya algo (aparte del desempeño de varias funciones) que necesite ser explicado. Pero si se toma la conciencia seriamente, las conclusiones a favor de las que estoy argumentando se siguen. (Chalmers 1996: 110)

Las intuiciones, de alguna manera, tienen que guiar la reflexión filosófica. De otro modo, la filosofía no tendría objeto; sin embargo, no está claro que planteen un dilema tan drástico a la reflexión. Por una parte, no habría cuestiones filosóficas interesantes si no hubiera diversas intuiciones en conflicto; en tal caso, si se espera algún tipo de progreso en la comprensión de un problema, probablemente no todas las intuiciones permanecerán intactas. Por otra parte, la noción de intuición no goza de un estatuto muy determinado; las intuiciones provienen de muy diversas fuentes a las que, en algunos casos, no tenemos un acceso transparente. Incluso cierta opacidad respecto a su origen es, precisamente, lo que posibilita que sean percibidas como datos básicos. En el debate sobre el problema de la conciencia, suelen aparecer de un modo recurrente argumentos que, en última instancia, remiten al carácter irrenunciable de una u otra intuición como un modo de rechazar una objeción. Sin embargo, estos argumentos, en la medida en que gocen de justificación, no pueden ser aplicados de un modo que excluya de antemano cualquier posibilidad de escrutinio o crítica. Un argumento que trate de mostrar que la concepción de la experiencia consciente que genera los problemas discutidos es incoherente no puede ser rechazado aduciendo que no toma la experiencia consciente en serio, o que niega las intuiciones de partida.

El rechazo contundente por parte de Chalmers de la crítica que Dennett lleva a cabo de cierta manera de entender la experiencia consciente obvia, como mostraré a continuación, la pertinencia de argumentos y de algunas intuiciones adicionales sustentados en la investigación empírica y, por otra parte, explota retóricamente cierta ambigüedad respecto a las connotaciones que puede suscitar la noción de “eliminación”, como un giro argumentativo que simplemente deja de lado el *explanandum* central. Sin embargo, ésta no puede ser toda la historia, ya que ignorar un *explanandum* no convierte a quien lo hace en alguien que se posiciona respecto al mismo, del mismo modo que un físico no es un eliminacionista

respecto a la biología por el hecho de no tener en cuenta el tipo de problemas que ésta trata de resolver.

Parte de los problemas que surgen en estas discusiones tienen que ver con cierta confusión entre cuestiones planteadas en el seno de nuestras prácticas cotidianas y cuestiones metafísicas; por supuesto, llevar a cabo esta distinción no presupone que el planteamiento de cuestiones metafísicas sea ilegítimo, ni que haya que dar prioridad a las posiciones de sentido común por defecto, ya que quizá la metafísica podría ser, en cierto sentido, una prolongación legítima de nuestras prácticas cotidianas y quizá nuestras prácticas cotidianas estén fundamentadas en ciertos supuestos metafísicos. La cuestión es que, en muchas ocasiones, el intento de abstraer ciertos conceptos de sus condiciones concretas de aplicación da lugar a categorías y prácticas que, superficialmente, pueden seguir reconociéndose como una extensión o perfeccionamiento de éstas, cuando realmente se han dejado de dar las condiciones bajo las cuales tiene sentido su aplicación, generando una ilusión conceptual. De un modo general, ésta es la estructura del tipo de crítica que presentamos en las siguientes secciones.

5.2. Realismo y familiaridad

La lógica por la que nacen y se consolidan los debates filosóficos en la actualidad se caracteriza, en ocasiones, por la rápida aceptación de una determinada manera de plantear los términos básicos de la discusión, y por el desarrollo posterior de una compleja dialéctica que gira alrededor de unas pocas intuiciones básicas que definen, en sus líneas generales y dependiendo de dónde se sitúe el énfasis, las distintas posiciones adoptadas. Más adelante, puede ocurrir que los análisis y las construcciones teóricas que pretendían aclarar y profundizar en la comprensión de las intuiciones de partida sean más oscuros que dichas intuiciones, situando la discusión en un callejón sin salida. Sin embargo, el peso de la discusión podría estar ocultando la genuina fuente de perplejidad. Pensemos, por ejemplo, en el planteamiento del funcionalismo. El supuesto de que los estados mentales conscientes son idénticos a estados funcionales, aun si se trata de un principio

heurístico que pretenda guiar una investigación, determinará tanto el tipo de debate que se llevará a cabo, como una investigación empírica definida. En el caso de Chalmers, incluso el tipo de posición anti-reduccionista que defiende se encuentra ligado en gran medida, aunque sea de un modo negativo, al espacio de investigación definido por la concepción funcionalista de la mente, de manera que lo fenoménico es el aspecto residual de la mente que no logra fijarse en el seno de una red causal.⁹⁰

Hasta ahora, he presentado algunos rasgos generales atribuidos a la experiencia consciente por Chalmers. La experiencia consciente goza de un carácter intrínseco, de manera que no puede individuarse de un modo relacional, y es accesible, exclusivamente, desde la primera persona, de manera que cada sujeto sólo tiene acceso a su propia conciencia. Si pensamos, por ejemplo, en el experimento del espectro invertido, este rasgo queda ilustrado claramente en la medida en que, dadas dos situaciones posibles, idénticas en todas sus propiedades físicas, neurológicas, psicológicas, etc., las propiedades fenoménicas de la experiencia cromática pueden variar. Por otra parte, el tipo de relación que el sujeto mantiene con sus propios estados es, según Chalmers, de familiaridad (*acquaintance*); Chalmers no precisa de un modo claro en qué pueda consistir esta relación pero, en cualquier caso, no se puede reducir a un vínculo causal, en la medida en que se trataría de un acceso mediado y contingente; en tal caso, los argumentos escépticos tradicionales podrían generarse, igualmente, respecto a la experiencia consciente, pero ésta es precisamente una conclusión que Chalmers tiene que rechazar, ya que, ¿qué podría significar que S tiene la experiencia fenoménica X, pero que posiblemente S se equivoca respecto al tipo de experiencia que tiene? ¿Sería una posibilidad conceptual que S tuviera una experiencia pero que le pareciera que no la tiene? Chalmers tiene que aceptar cierto tipo de

⁹⁰ En la historia de la filosofía, así como de la ciencia, hay ejemplos muy claros de cómo determinada concepción previa no sólo limitó la comprensión de determinadas cuestiones, sino que en algunos casos bloqueó completamente la investigación empírica. En este sentido, sería interesante tener en cuenta los debates a los que ha dado lugar, por ejemplo, la falta de precisión al definir distintos sentidos de causa, o la idea de que la existencia es un predicado. En el caso de la ciencia, piénsese, por ejemplo, en las posibilidades que abrió reconocer el principio de inercia o el cambio revolucionario en la comprensión de la noción de simultaneidad a partir de la teoría de la relatividad.

infallibilidad respecto a la propia experiencia consciente, en la medida en que la justificación de su posición se basa en el conocimiento en primera persona que cada sujeto tiene de su propia fenomenología. Su fuerte realismo respecto a las otras mentes le fuerza a comprometerse con el requisito de que el acceso a los propios estados mentales conscientes debe tener un carácter *sui generis*; de otro modo, ¿qué sentido tendría hablar de estados mentales conscientes a los que no sólo no se tiene acceso desde el punto de vista de la tercera persona, sino que resultan opacos para el mismo sujeto que los tiene?

Algunos filósofos apuntan a que, al menos, tenemos un acceso a nuestra propia experiencia que es más limitado de lo que presuponemos. Ahora bien, ¿cómo podría llegar a establecerse una conclusión de este tipo si, en última instancia, los testimonios desde el punto de vista de la primera persona son el tribunal último respecto a la conciencia? Como señalé más arriba, la transposición de determinadas estructuras conceptuales a situaciones en las que no tiene sentido su aplicación puede producirse de un modo inadvertido. En cierto sentido, la idea de una concepción absoluta de la realidad, de la que hablé en el Capítulo 1, es deudora de la idea de que el conocimiento se encuentra ligado, necesariamente, a una perspectiva, y de un modo negativo trata de eliminar los rasgos contingentes de la condición humana que dan lugar a una perspectiva sesgada de la realidad. En contextos bien definidos, la idea de que se tiene una perspectiva sesgada de la realidad tiene un sentido muy claro. Por ejemplo, es sencillo encontrar a diario descripciones periodísticas en el ámbito de la política que claramente dejan de lado información o datos que resultan relevantes para la evaluación de una determinada cuestión, y que en general favorecen un fin o punto de vista previo interesado en excluir dicha información; sin embargo, ¿resulta inteligible la idea de una perspectiva respecto a los asuntos políticos que no dependa, en absoluto, de un determinado interés, o que no se encuentre ligada a la situación particular de un grupo de personas? ¿Acaso la denuncia de una situación injusta no describe un hecho adecuadamente simplemente porque esta denuncia la lleva a cabo la misma víctima de la injusticia, y responde a un interés particular?

El tipo de argumentación que presento a continuación se basa en la idea de que, en los debates sobre la experiencia consciente, y contrariamente a lo que podría parecer, no se toma lo suficientemente en serio la idea de que la conciencia se encuentra conceptualmente ligada a unas condiciones empíricas definidas, obviando que se están dejando de lado las condiciones necesarias que hacen posible la inteligibilidad de lo que supone ser un sujeto que goza de experiencia consciente.

5.3. El materialismo cartesiano

Una idea poco problemática respecto a la experiencia consciente es que el tipo de experiencias que tenemos se encuentra determinado por el momento y el lugar en el que nos encontramos. Una condición necesaria para que haya experiencia consciente es la existencia de estímulos sensoriales. Si no hay luz que llegue a nuestras retinas, no es posible tener experiencias visuales, y si no hay ondas sonoras que hagan vibrar nuestros tímpanos o una excitación de las células nerviosas de nuestra piel, no tendremos experiencias sonoras ni táctiles, respectivamente.⁹¹ Una ilustración clara de cómo el tipo de experiencias conscientes que tenemos depende de nuestra situación espaciotemporal es la contemplación de un espectáculo de fuegos artificiales. Cuando observamos en el cielo la explosión de un artefacto pirotécnico, transcurre un pequeño lapso de tiempo hasta que oímos su explosión. Como es sabido, esto se debe a las diferentes velocidades de la luz y el sonido: el desfase temporal entre ambas sensaciones se explica, simplemente, porque las ondas sonoras se demoran más en llegar al punto en el que nos encontramos. Igualmente, dependiendo de en qué posición nos

⁹¹ Estrictamente, esto no es verdad, en la medida en que estos tipos de sensaciones pueden ser producidos de manera anómala, o en circunstancias que no tienen que ver con la percepción del mundo externo. Por ejemplo, pensemos en las imágenes geométricas que se forman con los ojos cerrados, o la molesta sensación sonora llamada *tinnitus*, que no se corresponde con ninguna estimulación externa. Sin embargo, que la fuente de estimulación no provenga de una fuente externa al cuerpo no implica que no haya una estimulación sensorial, ni que ésta no tenga una localización espacial determinada. La idea de lo externo es notoriamente ambigua, y si se reduce a distinguir lo que se encuentra debajo de la piel de lo que se encuentra fuera, no es demasiado útil para comprender la experiencia consciente.

situemos respecto a un objeto, tendremos distintas experiencias visuales del mismo, debido a la incidencia de la luz en la retina.

La secuencia compuesta por la estimulación sensorial, la experiencia consciente y una posible reacción posterior por parte del sujeto (una preferencia verbal, un movimiento corporal, etc.) parece natural e incuestionable. Sin embargo, ¿en qué medida puede este principio ser aplicado a casos en los que hay implicadas magnitudes temporales ínfimas? ¿Hay intervalos de tiempo de duración determinada para la experiencia consciente? Respecto al espacio, ¿puede decirse que hay una línea que deba ser traspasada por la señal nerviosa para que haya experiencia consciente? Es evidente que la formación de una imagen en la retina es una condición necesaria de la visión, pero esto no es suficiente, ya que esta información tiene que transmitirse a través del nervio óptico. La cuestión es, ¿hay un punto del cerebro hasta el que pueda rastrearse dicha señal nerviosa que permita distinguir un recorrido pre-experiencial de un recorrido post-experiencial? Una concepción que se comprometa con una respuesta afirmativa a esta pregunta, un vez rechazado el dualismo cartesiano, es la posición que Daniel Dennett denomina materialismo cartesiano. Así es como lo describe en relación a un experimento en el que los participantes en el mismo tienen que presionar un botón al tener determinadas experiencias sensoriales:

[N]uestra intuición natural es que la experiencia de la luz o el sonido ocurre entre el tiempo en que las vibraciones golpean nuestros órganos sensoriales y el tiempo en que logramos apretar el botón señalando esa experiencia. Y ocurre en algún lugar centralmente, en algún lugar en el cerebro en los senderos estimulados entre el órgano sensorial y el dedo. Parece que, si pudiéramos decir exactamente dónde, podríamos decir exactamente dónde ocurrió la experiencia. Y viceversa: si pudiéramos decir dónde ocurrió exactamente, podríamos decir en qué lugar del cerebro se situó la experiencia.

Llamemos a la idea de tal locus central en el cerebro materialismo cartesiano, puesto que es la visión a la que llegas cuando rechazas el dualismo de Descartes, pero no

logras rechazar la imagen de un teatro central (aunque material) donde “todo se une”. (Dennett 1991: 107)⁹²

La idea de que hay un lugar y momento en el que, de algún modo, se presenta la conciencia, es lo que Dennett considera la concepción del Teatro Cartesiano. Esta definición recoge una intuición muy potente a la que es difícil sustraerse. De hecho, en cierto sentido, no puede sino ser verdadera, si atendemos a los ejemplos presentados. Incluso en relación con un estudio empírico de la conciencia, ¿de qué otro modo podría llevarse a cabo una investigación sobre ésta si no se presupone la existencia de coordenadas espaciotemporales que la delimiten? Por ejemplo, una distinción tan básica como la ya conocida desde hace siglos entre un movimiento voluntario y un reflejo parece presuponer que la señal nerviosa no llega, en ocasiones, a traspasar cierta línea, o a alcanzar cierto punto que, de un modo privilegiado, se vincula a la aparición de experiencia consciente. Sin embargo, la formulación de esta tesis, tal y como aparece en la cita, parece irrelevante por lo que concierne a la fenomenología. En la medida en que esta posición es un tipo de materialismo, no está claro que atienda al sentido de conciencia que Chalmers considera problemático. Dado que Chalmers es un dualista, es difícil que asienta a la idea de que la conciencia se da en un punto del cerebro. Lo que distingue a los zombis de seres humanos normales no es ninguna diferencia material, y la discusión en torno a la concepción del teatro cartesiano podría dirimirse sin que el problema difícil de la conciencia haya sido abordado; de hecho, en el ámbito de la ciencia cognitiva es difícil encontrar teorías que, de un modo general, se comprometan con la concepción del Teatro Cartesiano.⁹³ Hablar de un lugar en el que “ocurre” la experiencia parece dar por sentado que es el concepto psicológico de lo mental el que está en juego pero, ¿en qué sentido una experiencia fenoménica de rojo puede

⁹² Quizá el sentido en el que Dennett tiene en cuenta la idea de un lugar central en el que se da la conciencia no coincide adecuadamente con una tesis que haya sido sostenida explícitamente por Descartes. La función que Descartes atribuye a la glándula pineal es el de un lugar de interacción entre la *res cogitans* y la *res extensa* y no un lugar en el que se da la experiencia consciente. Como veremos a continuación, la definición de Dennett se puede modificar de manera que atrape mejor la tesis que realmente está en juego.

⁹³ Piénsese en la Global Work Space Theory de Baars (1988) e incluso en la hipótesis de la oscilación de Crick (1994).

darse en el cerebro? ¿Cómo es posible, por ejemplo, que nuestras experiencias cromáticas se encuentren en el cerebro cuando, de hecho, no hay colores en el cerebro (al margen de una monótona tonalidad que poco tiene que ver con la riqueza de nuestra experiencia)?

No obstante, no hay que olvidar que el dualismo es una teoría que trata de dar cuenta de los hechos más cotidianos sobre la experiencia consciente e incluso, en el caso de la posición Chalmers, hay un intento de fundamentar una teoría científica de la experiencia consciente. En este sentido, negar la tesis general del materialismo no implica renunciar a la idea de que la dimensión de lo mental está exenta de regularidades y de conexiones uniformes. Como vimos, Chalmers no niega la supervinencia *tout court* de la experiencia consciente sobre los hechos físicos, sino sólo la supervinencia lógica; la diferencia entre el materialismo y el dualismo de Chalmers se reduce, de este modo, a una cuestión modal que, en términos prácticos, es sutil. En última instancia, el hecho de que la experiencia consciente no sea reducible a hechos físicos no es algo más problemático que la postulación de propiedades electromagnéticas que en su momento la física tuvo que llevar a cabo, al tratarse de un fenómeno no reducible a hechos mecánicos. Por otra parte, hay dificultades profundas respecto al tipo de ciencia de la conciencia que propugna Chalmers. Estas dificultades serán tratadas con mayor detalle en el próximo capítulo. Pero hay un rasgo del aspecto fenoménico de lo mental relevante para la presente discusión que tiene que ver con el modo en que la fenomenología se encuentra anclada al mundo. Hay un sentido en el que, una vez aceptado el dualismo respecto a lo mental, deja de ser inteligible la idea de que hay un lugar en el que se da la conciencia, a modo de teatro cartesiano. Recordemos que el problema que plantea la experiencia fenoménica no quedaría resuelto, aunque encontrásemos en el cerebro pequeñas réplicas de las cosas y propiedades que se perciben, ya que, de todos modos, el acceso a cómo se presentan esas cualidades al sujeto nos seguiría estando vedado. Sin embargo, la teoría de Chalmers vincula la experiencia consciente a un lugar privilegiado de un modo decisivo. En el siguiente fragmento, podemos encontrar una ilustración del tipo de posición que defiende Chalmers:

No obstante, resulta todavía posible que la conciencia *surja* de una base física, incluso aunque no esté implicada por dicha base. La posición con la que quedamos es que la conciencia surge de un sustrato físico en virtud de ciertas leyes contingentes de la naturaleza que no están implicadas en sí mismas por leyes físicas. Esta posición es sostenida implícitamente por mucha gente que se considera a sí misma materialista. Es habitual oír “por supuesto que soy un materialista; la mente surge, ciertamente, del cerebro”. La misma presencia de la palabra “surge” debería considerarse reveladora. No se suele decir “aprender surge del cerebro”, por ejemplo—y si se hiciera sería en un sentido temporal de “surge”. Más bien, sería más natural decir que aprender es un proceso en el cerebro. El mismo hecho de que la mente necesite *surgir* del cerebro indica que hay algo adicional que está ocurriendo, más allá de los hechos físicos. (Chalmers 1996: 125-26)

Según Chalmers, la experiencia consciente ha de relacionarse espacialmente con el cerebro. De acuerdo con este sentido mínimo de vínculo espacial, la experiencia consciente *surge*, de algún modo, del cerebro, y no constituye un proceso, en la medida en que lo pueda ser una disposición psicológica como el aprendizaje. Por otra parte, en contraposición a la idea de que el aprendizaje surge del cerebro en un sentido temporal, ¿cuál es el otro sentido de surgir según el cual la conciencia surge del cerebro? Un problema relacionado con el intento de anclar la fenomenología al mundo después de haber llevado a cabo un compromiso con una concepción intrínseca de la conciencia, y por ello divorciada conceptualmente del cuerpo, es que no es posible volver a dotar de inteligibilidad a la idea de que la experiencia consciente se encuentra ligada de un modo especial a un *locus*. Por una parte, aun si fuera posible individuar estados mentales conscientes con independencia de ser adscritos a un cuerpo y sin ninguna relación con el mundo circundante,⁹⁴ habría que resolver la dificultad adicional de determinar en virtud de qué puede decirse que un determinado estado mental surge de un conjunto de hechos físicos y no de otro, ya que la experiencia consciente no aparece en ningún lugar, a no ser en un lugar metafórico. Podría recurrirse a la sucesión regular entre

⁹⁴ En el Capítulo 1 traté de mostrar que esta forma de concebir la experiencia consciente es incoherente.

la instanciación de cierto tipo de hechos físicos⁹⁵ y la instanciación de propiedades fenoménicas. Pero, en tal caso, ¿por qué considerar que la experiencia consciente surge de un cerebro particular? Probablemente, siempre que mis fibras C disparan al sufrir un traumatismo, las fibras C de otra persona están disparando; ¿por qué no considerar que cada vez que tengo dolor esta experiencia surge del sistema nervioso de una persona distinta, de las muchas que en ese momento ejemplifican la propiedad funcional correlacionada con el dolor fenoménico? ¿O qué razón independiente tendríamos, por ejemplo, para descartar la disyunción de todas las instanciaciones de propiedades funcionales de dolor como el correlato de *mi* dolor fenoménico, surgiendo este último de dicha disyunción?

Estas implicaciones, dejando de lado su implausibilidad, nos pueden ayudar a comprender mejor en qué medida la caracterización de Dennett del Teatro Cartesiano recoge una manera dominante de concebir la experiencia consciente en los debates actuales. En realidad, el compromiso con que la experiencia se dé en un lugar concreto del cerebro no es una idea tan plausible como pudiera parecer inicialmente, y no goza de gran aceptación en los debates actuales. Por una parte, incluso en la investigación empírica esta imagen ya ha dejado de ser dominante, de manera que el tipo de modelos explicativos que se tienen en cuenta no se comprometen con la existencia de un punto privilegiado del cerebro, sino que más bien tratan de comprender la interacción entre distintas partes del cerebro y, en algunos casos, el cuerpo, de manera que la conciencia no puede llegar a entenderse privilegiando una determinada región del cerebro. Por otra parte, un dualismo como el de Chalmers, que en última instancia se diferencia sutilmente de una posición materialista, parece quedar fuera del alcance de la concepción del Teatro Cartesiano, en la medida en que la caracterización de lo mental que ofrece excluye la vinculación conceptual de la fenomenología a un *locus*.

Sin embargo, hay cierto sentido según el cual la idea de un lugar o lugares privilegiados sigue teniendo importancia. Si las preguntas que se plantean respecto

⁹⁵ Uno de los problemas que se presentaban al tratar de establecer una correlación entre propiedades funcionales realizadas físicamente y estados mentales conscientes es que no está claro qué puede contar, previamente a la identificación del tipo de experiencia consciente en cuestión, como un patrón funcional relevante que sea correlato de la misma.

a la experiencia consciente incluyen términos como “correlatos de la experiencia”, “surgimiento de la experiencia”, o si simplemente se plantea si dada cierta propiedad funcional hay una propiedad fenoménica que, en cierto sentido, la *acompaña*, parece que la experiencia consciente se está concibiendo, por decirlo de algún modo, de acuerdo con un formato metafísico definido, que pueda situar en paralelo la fenomenología con propiedades y hechos no fenoménicos. Si el dolor funcional, por ejemplo, es el disparo de las fibras-C causado por un daño de tejido, entonces hay una propiedad fenoménica correspondiente que, si bien no se da en espacio alguno, se vincula, de un modo específico, con determinado *locus*. Si es posible establecer una correlación entre estados funcionales y estados mentales conscientes, la conciencia heredará, de un modo general, una estructura metafísica isomórfica con un conjunto relevante de estados mentales psicológicos (funcionales).⁹⁶ En este sentido, la experiencia consciente puede ser fragmentada de un modo discreto, de manera que una experiencia de rojo tiene que poder corresponder, aunque sea de un modo abstracto o lógico, con las coordenadas espaciotemporales de la propiedad funcional de la que se supone que surge. No obstante, ¿no tenemos la impresión de que la experiencia consciente no puede fragmentarse de esta manera? ¿Acaso la conciencia no es un flujo continuo y constante?

Una concepción discreta de la experiencia, entendida en un sentido vago, resulta inocua. En muchos casos, nuestra manera de concebir las sensaciones es claramente discreta. Pensemos, por ejemplo, en cómo comparamos los sabores de distintos vinos: de algún modo, centramos nuestra atención en la sensación gustativa y tratamos de retener en la memoria su carácter específico para poder ser comparado, posteriormente, con el sabor de la otra muestra de vino. En cierta manera, retener el carácter de la experiencia gustativa implica mantenerla aislada de fuentes que la puedan contaminar, y para ello se toman precauciones físicas, como limpiar la copa antes de servir la segunda muestra o enjuagar la boca, pero también se intenta que el recuerdo de otros sabores, o incluso otras sensaciones,

⁹⁶ Ésta es, expresada de una manera burda, la concepción general de los principios de la ciencia de la conciencia propuestos por Chalmers, de la que más adelante hablaremos.

no se interpongan y, para ello, es importante *prestar atención a la sensación*. Ahora bien, ¿cómo debe entenderse la idea de que podemos prestar más o menos atención a las sensaciones? ¿Puede decirse que las sensaciones están presentes, como lo puede estar un objeto en una habitación, y pasar *desapercibidas* o ser ignoradas? ¿O bien el acto de prestar atención es constitutivo de la sensación? Parece que en un sentido poco reflexivo no hay ningún problema en la idea de que podemos no prestar atención a una sensación, a pesar de que ésta se dé. Al fin y al cabo, hay una clara diferencia entre no prestar atención a un sonido, de manera que pasa desapercibido, y estar sordo durante el momento en el que se produce el mismo. Sin embargo, ¿cómo se combina esta posibilidad con el principio de Chalmers de que, respecto a la experiencia consciente, no se puede trazar una distinción entre ser y parecer?

Si se consideran inteligibles los argumentos de los que se sirve Chalmers, en ellos se presupone que, de un modo general, la pregunta por la instanciación de una determinada propiedad fenoménica tiene, en cualquier caso, una respuesta definida. En el caso de la persona cuyo espectro cromático se encuentra invertido, no hay una manera definitiva de saber si de hecho tal es el caso, ni de determinar cuál es el color, en el sentido fenoménico, que dicha persona percibe.⁹⁷ Sin embargo, el experimento presupone que hay una *única respuesta verdadera* a la pregunta de si el espectro cromático de dicha persona se encuentra invertido respecto al nuestro, aunque no podamos conocerla con certeza, ni siquiera en condiciones epistémicas ideales. Este *requisito fuertemente realista*, unido a la idea de que la experiencia consciente adopta un *formato metafísico discreto*, es difícil de mantener coherentemente junto con la idea de que nuestro acceso a la experiencia consciente puede explicarse apelando a la relación de familiaridad con un tipo de propiedad fenoménica, como veremos en el siguiente apartado.

He presentado la concepción del Teatro Cartesiano de Dennett como el principio de que hay unas coordenadas espaciotemporales en las que se da la

⁹⁷ Obsérvese que en el caso de la cata de vinos este tipo de cuestión no se plantea, de manera que, aunque se puedan aislar sabores con una gran precisión, en ningún momento se plantea la posibilidad de que ante dos discriminaciones coincidentes respecto a un mismo vino, llevadas a cabo por dos personas distintas, la sensación gustativa sea radicalmente distinta.

conciencia. Si bien tal y como se formula no parece expresar un supuesto necesario para formular el problema difícil de la conciencia, apunta en la dirección adecuada en la medida en que cualquier teoría actual de la conciencia que tome en serio el papel del cerebro como la base de la que surge la conciencia, y que trate de respetar nuestras intuiciones básicas respecto a la identidad de los estados mentales, de algún modo u otro tiene que vincular la fenomenología con un locus, aunque sea de un modo abstracto. A su vez, la inteligibilidad de los argumentos presentados por Chalmers exige que se presuponga cierta forma metafísica para la experiencia consciente, según la cual las distintas cualidades fenoménicas gozan de un *formato metafísico discreto* al que se añade una *concepción fuertemente realista*. A continuación, presentaré una argumentación que tratará de mostrar la ininteligibilidad de estos dos últimos supuestos.

5.4. La crítica del Teatro Cartesiano

En la sección anterior, sugerí que la confusión entre cuestiones planteadas en el seno de nuestras prácticas cotidianas y cuestiones metafísicas se encontraba detrás de las perplejidades filosóficas planteadas a propósito de la experiencia consciente. En esta sección, revisaré el análisis de Dennett del fenómeno phi que, tal y como mostraré, apuntan a la ininteligibilidad de los supuestos metafísicos que posibilitan la formulación del problema difícil de la conciencia.

5.4.1. Revisiones estalinistas y revisiones orwellianas

Uno de los casos centrales de los que se sirve Dennett en *Consciousness Explained* (1991) para rebatir la concepción del Teatro Cartesiano es el llamado fenómeno phi. Dicho fenómeno es el que ha permitido el desarrollo del cine y la televisión, de modo que una serie de fotogramas no se percibe como una sucesión de instantáneas, sino como movimiento.⁹⁸ Una forma de generar esta ilusión consiste en proyectar

⁹⁸ En Wertheimer (1912) aparece la investigación pionera en este tipo de fenómeno, aun cuando ya existía el cine. El tratamiento del fenómeno tal y como aquí lo presentamos se remite a los estudios de Kolers y Grünau (1976).

dos focos luminosos en rápida sucesión y en un ángulo visual de cuatro grados, de manera que el sujeto percibe un único foco luminoso desplazándose de izquierda a derecha. Una modificación del experimento, sugerida a Kolers por Nelson Goodman⁹⁹, consistió en proyectar focos luminosos con distinto color. El resultado del experimento fue, de un modo imprevisible, la percepción, por parte de los sujetos, de un foco luminoso en movimiento que cambia abruptamente de color en la mitad de su recorrido. Ahora bien, ¿cómo se explica la aparente anticipación por parte del sujeto del que va a ser el color del segundo foco proyectado?¹⁰⁰

Aceptando el principio de que la experiencia consciente se vincula discretamente a determinado momento y lugar del cerebro, y excluyendo, por supuesto, que en los experimentos se produzca algún tipo de anticipación del futuro, de manera que el fenómeno sólo puede producirse una vez percibida la proyección del segundo foco, Dennett distingue dos posibles teorías que dan cuenta del fenómeno. De acuerdo con la primera de ellas, hay una experiencia consciente del segundo punto luminoso pero, rápidamente, el cerebro, de algún modo, interpreta la información que recibe y modifica el recuerdo de la experiencia incorporando la percepción de las fases sucesivas que componen el desplazamiento de un único foco entre los dos puntos; éste es un tipo de teoría orwelliana, recordando la célebre novela *1984*, en la que el Ministerio de Verdad altera la versión sobre el pasado histórico en función de los intereses estatales coyunturales. La segunda teoría tiene que introducir un elemento teórico adicional. Según ésta, no llega a haber una experiencia consciente previa que luego será borrada de la memoria, sino que se produce un retraso en la aparición de la conciencia de manera que, a modo de una censura televisiva, una vez se ha producido una percepción subpersonal de los dos focos, la experiencia consciente “proyectada” es un montaje ilusorio, en el que se introducen las fases intermedias que componen el movimiento del foco; esto es lo

⁹⁹ Véase Goodman (1978).

¹⁰⁰ En el caso del experimento inicial igualmente se produce una aparente predicción del segundo estímulo, en la medida en que se percibe la dirección de la trayectoria de acuerdo al lugar en el que se va a proyectar el segundo foco, aunque como reconoce Dennett, en la variación del experimento se aprecia de un modo más claro este fenómeno.

que Dennett llama una teoría estalinista de la conciencia, en alusión a los juicios montaje perpetrados durante el estalinismo.

Lo que diferencia a la teoría orwelliana de la estalinista es que, en el caso de la primera, se postula un momento en el que hay una experiencia consciente que no permanece, posteriormente, en el registro de la memoria, al ser sustituido por un nuevo registro, mientras que la teoría estalinista se compromete con un retraso respecto al momento que se produce la experiencia que pasa directamente a la memoria. Dicho de otra manera, en el caso de la teoría orwelliana se produce una ilusión memorística, mientras que en el caso del montaje estalinista se produce una alucinación visual. Si aceptamos la posición realista respecto a la experiencia consciente que hemos descrito más arriba, al menos una de las dos teorías debería ser verdadera.¹⁰¹ Ahora bien, ¿qué podría contar como evidencia para determinar cuál de las teorías es la verdadera?

Hasta ahora, no se ha dicho nada respecto a cuál pueda ser el criterio que nos permita determinar si en un momento dado ha habido conciencia o no. Ya vimos que en contextos prácticos cotidianos hay criterios muy claros de cuándo una persona es consciente de un estímulo, como una constatación verbal o una reacción corporal. En el caso de los experimentos sobre el fenómeno phi, de entrada, el único criterio para determinar el momento de la conciencia son los informes verbales llevados a cabo por el sujeto. Hay otros medios posibles, como el uso de un botón, que en principio podrían fijar con mayor precisión el momento en el que se produce la experiencia consciente. Ahora bien, en este caso, ¿podrían servir dichos criterios para decidir cuál de las dos teorías es la verdadera? La teoría estalinista tiene que afrontar una dificultad importante: al postular un retraso de tiempo, durante el cual se produce el “montaje”, debería, en general, apreciarse determinado lapso de tiempo fijo constante para la percepción consciente de cualquier estímulo; sin embargo, según Dennett, el lapso de tiempo implicado, 200 milésimas de segundo,

¹⁰¹ El espacio lógico de posibles teorías que den cuenta del fenómeno no se agota en estas dos teorías: Por ejemplo, podríamos pensar en una teoría barroca en la que hay varios registros y borrados en la memoria, o una especie de combinación entre una teoría estalinista y una orwelliana, de modo que hubiera tanto un montaje como el borrado del registro de un recuerdo anterior. Estas posibilidades quizá requerirían unos intervalos de tiempo inverosímiles. En cualquier caso, el hecho de que el espacio lógico de posibles teorías sea más amplio es irrelevante para la argumentación, como mostraré a continuación.

es demasiado amplio dados los estudios que hay respecto a respuestas bajo control consciente, de modo que el botón se apretaría antes de percibir el segundo foco.¹⁰² No obstante, el teórico estalinista todavía podría argüir que el botón se presiona inconscientemente, de manera que el estímulo del primer foco se procesaría, inicialmente, de un modo subpersonal. La cuestión es, si aceptamos la idea de que la experiencia consciente se puede fragmentar discretamente, ¿hay algún modo de determinar si una de las dos teorías es verdadera? La respuesta de Dennett es que no.

Por una parte, no hay ningún criterio desde el punto de vista de la tercera persona que pueda decidir entre ambas teorías. Un informe verbal constituye, en circunstancias normales, un criterio habitual para atribuir conciencia; pero en este caso, la escala de tiempo implicada es demasiado pequeña para que este criterio permita establecer alguna diferencia; igualmente, respecto al ejemplo imaginario de la presión de un botón hemos visto que se pueden proponer dos alternativas incompatibles que dicho criterio no puede discriminar. Esta limitación no debería sorprendernos si se acepta que el acceso a la experiencia consciente desde la tercera persona es falible, en un sentido en el que no lo es, por ejemplo, el conocimiento de objetos externos. De hecho, ningún indicio desde el punto de vista de la tercera persona podría ayudarnos a distinguir un ser humano normal de un zombi. Ahora bien, ¿hay alguna manera de discriminar entre ambas teorías desde el punto de vista de la primera persona? Precisamente, lo que motiva la propuesta de dos teorías rivales es un único testimonio que describe un foco desplazándose de izquierda a derecha, y que cambia de color en la mitad del recorrido; ninguna de las manifestaciones públicas de los sujetos del experimento, incluidas las manifestaciones verbales, pueden dirimir la cuestión. Por otra parte, apelar a la

¹⁰² En este punto, Dennett no es demasiado claro. Precisamente, se está poniendo en cuestión la idea de que puede determinarse cuándo se produce, exactamente, el momento de la conciencia; sin embargo, esta observación parece presuponer que, de hecho, dicho momento se puede determinar. Ésta no es una dificultad demasiado importante para Dennett, si se entiende que Dennett no niega que haya un lapso de tiempo entre un estímulo y la percepción consciente del mismo, sino la existencia de una línea bien definida que deje a un lado una senda preexperiencial, con un estímulo no “consciente”, y una senda posexperiencial, de manera que siempre haya una respuesta definida respecto a si se ha instanciado una determinada propiedad fenoménica. Este último tipo de cuestión es muy distinto del intento de determinar, por ejemplo, el transcurso de tiempo entre un estímulo y un informe verbal.

familiaridad que cada sujeto tenga con su experiencia, al margen de que la subsiguiente expresión verbal sea igualmente predicha por ambas teorías, no será de gran ayuda. Un movimiento en esta dirección llevaría a afirmar que hay un sentido en el que el sujeto tiene una experiencia fenoménica que no se corresponde con su posterior informe verbal respecto al tipo de experiencia que le *parece* tener. Pero no olvidemos que la argumentación de Chalmers suponía que, respecto a la experiencia consciente, no se puede trazar una diferencia entre ser y parecer, a no ser que se admita la categoría de lo “objetivamente subjetivo”, o como Dennett apunta recordando el diálogo de Raimond Smullyan, “el modo en que las cosas te parecen real y objetivamente, incluso si no parecen parecerte de ese modo” (Smullyan 1981).

Independientemente de que resulte contraintuitivo, la aceptación de este principio conduce a tres consecuencias inaceptables. En primer lugar, aunque puede sostenerse que hay cierta ambigüedad respecto al término “parecer” (de manera que según cierto sentido de “parecer” respecto a la experiencia consciente no se puede diferenciar su ser de su parecer, como en el caso de un objeto físico, pero según otro sentido de “parecer”, quizá de un orden distinto, nuestra experiencia consciente puede parecerse distinta de lo que es) esta distinción convertiría en irrelevante el supuesto rasgo exclusivo de la experiencia consciente que Chalmers presupone para poder aplicar la herramienta bidimensional, de manera que ya no podría decirse de un modo claro que las intenciones primarias son las relevantes al evaluar las posibilidades de los zombis o el espectro invertido. En segundo lugar, tornaría en inútil cualquier apelación a un conocimiento inmediato de la propia experiencia consciente como una especie de tribunal último, de manera que las intuiciones que constituían el punto de partida en la reflexión sobre la fenomenología dejarían de ser fiables. Finalmente, y esta es la dificultad más grave a la que llevaría el planteamiento de lo objetivamente subjetivo, el supuesto estado de cosas que dirimiría el conflicto entre las dos teorías es ininteligible, en la medida en que éste no establece ninguna diferencia sustantiva, tanto desde el punto de vista de la tercera persona como del de la primera persona o, como señala Dennett, “Esta es una distinción que no establece ninguna distinción” (Dennett 1991: 125). Esta última

conclusión es la que ha llevado a varios autores a considerar que el planteamiento de Dennett presupone algún tipo de verificacionismo. A continuación, argumentaré que esta objeción es infundada y trataré de mostrar que la noción de fenomenología que critica Dennett es ininteligible en un sentido independiente de la verdad de una teoría verificacionista.

5.4.2. El Teatro Cartesiano y el verificacionismo

La crítica del Teatro Cartesiano que Dennett lleva a cabo sirviéndose de la reflexión sobre el fenómeno phi se resume en la idea de que no hay un hecho o conjunto de hechos que determine la diferencia entre una senda preexperiencial y una senda posexperiencial, de manera que la experiencia consciente no puede vincularse a un lugar ni un momento determinados en toda situación posible, como presupone la argumentación de Chalmers. Ahora bien, ¿no es ésta una forma de verificacionismo que, en cualquier caso, requeriría de una justificación adicional? La respuesta a esta pregunta dependerá, por supuesto, de qué se entienda por verificacionismo y en qué medida la argumentación de Dennett se ajuste al modelo de verificacionismo propuesto. Kevin B. Korb plantea una objeción en este sentido:

Es simplemente falso que los dos modelos sean, simplemente, “verbalmente diferentes” el uno del otro: la confusión de una incapacidad para probar una teoría frente a la otra *ahora* con la incapacidad *en principio* para probar ambas es una confusión fundamental. Es lo mismo que confundir el operacionalismo con la exigencia de que nuevas teorías deban hacer *alguna* posible *diferencia* en nuestras observaciones, que es un prerrequisito de que haya *alguna ciencia experimental posible* que decida entre esas teorías. (Korb 1993: 5)

En la observación de Korb se distinguen dos sentidos diferentes en los que la verdad o falsedad de una teoría no puede ser probada. Uno de ellos alude a una situación efectiva que, de un modo contingente, no permite decidir entre dos teorías; un ejemplo claro de este caso es la limitación tecnológica de un determinado momento, aunque también podría incluir casos en los que es imposible, incluso en un futuro,

establecer la verdad de una teoría o proposición, por ejemplo, en referencia a hechos pasados de los que no se dispone evidencia. En cualquier caso, el contraste con el segundo sentido es muy claro, aludiendo este último a una imposibilidad, *en principio*, de dirimir entre dos teorías. Es precisamente la idea de que dos teorías predicen distintas observaciones lo que hace posible la ciencia empírica, pero este principio básico de sustantividad mínima de una teoría no debe ser confundido con la simple imposibilidad fáctica de decidir entre dos teorías, o con una teoría reduccionista de los términos teóricos que identifique su significado con un determinado método de verificación. Ahora bien, ¿confunde Dennett ambos sentidos de imposibilidad en su argumentación respecto al fenómeno phi? La cuestión es, ¿qué podría contar como un criterio concluyente?

Korb sugiere que, si bien una manifestación verbal de conciencia dura demasiado tiempo para resultar relevante, quizá haya estados cerebrales que se produzcan en un intervalo de tiempo muy corto y que se correlacionen, en situaciones normales, con dichas manifestaciones; la presencia o ausencia de estos fugaces estados cerebrales podría marcar la diferencia, aunque hoy en día no la podamos detectar. Sin embargo, ¿cómo podría establecerse esta correlación en primera instancia? Dependiendo del momento concreto en el que se produzca la correlación entre ambos sucesos, ¿en virtud de qué puede determinarse que dicho estado cerebral señala el momento de la conciencia? La argumentación de Dennett no apela, en ningún momento, a las limitaciones actuales de la tecnología o a algún otro tipo de limitación contingente, sino que trata de mostrar cómo determinada concepción de la experiencia consciente formada a partir de unas condiciones de aplicación determinadas no puede extrapolarse inteligiblemente como un modelo de comprensión último de la misma y, aun si hubiéramos llegado a superar toda limitación epistemológica, la cuestión respecto al momento en que se produce experiencia consciente permanecería indeterminada.

5.4.3. Escepticismo y sustantividad

El tipo de confusión al que apunta Korb no es especialmente precisa, en la medida en que la idea de que no se pueda decidir, *en principio*, entre dos teorías es demasiado ambigua. William Seager (1999) considera que el modo del que Dennett se sirve del verificacionismo es asimilable al planteamiento de un escenario escéptico, proponiendo un símil que permite una comparación más precisa:

Esto es equivalente a la hipótesis escéptica de que no puedes *saber* si eres consciente o no (incluso concediendo que haya una cuestión de hecho), puesto que ésta es una afirmación respecto al pasado (muy) reciente y podría haber una implantación constante de un recuerdo falso de la creencia de que eres consciente. Aunque tan intrigante y enigmático como todos los ataques escépticos al conocimiento, la versión verificacionista de Dennett de la posición de que no hay conciencia no es más aceptable que la escéptica. (Seager 1999: 127)

En este caso, si no es posible distinguir entre dos hipótesis no es porque haya una situación contingente o una limitación empírica, sino que no es posible imaginar una situación en la que se pueda probar una de las dos hipótesis. No obstante, al servirse de esta comparación, Seager confunde la finalidad de la argumentación de Dennett. Un ejemplo paralelo de argumentación del que Seager se ha servido previamente para desacreditar el verificacionismo es la famosa hipótesis russelliana de que el mundo ha sido creado hace cinco minutos. Tanto la argumentación escéptica relativa a la implantación continua de una memoria falsa, como la de la creación reciente del mundo comparten un rasgo entre sí que las diferencia de la argumentación de Dennett. En el caso presentado en la cita, no se pretende mostrar que determinada concepción de la experiencia consciente es ininteligible, sino que la misma formulación del escenario presupone la comprensión de lo que significa ser consciente, en contraste con haber tenido el recuerdo de haber sido consciente; del mismo modo, la hipótesis de la creación reciente no trata de mostrar que no tiene sentido la idea de un mundo que existe desde hace más de cinco minutos, sino que ambas hipótesis son inteligibles y compatibles con toda la evidencia de la que disponemos. En el caso de la teoría estalinista y la teoría orwelliana, todavía no se ha dado el paso previo que permite generar un argumento escéptico, dado que

ni siquiera hay una inteligibilidad de las hipótesis en conflicto. Y éste es el resultado, como ya he señalado, de extrapolar una determinada concepción de la experiencia consciente a casos en los que ésta no puede aplicarse con sentido. Si el tipo de pregunta que formulamos tiene que ver con si una persona es consciente de un estímulo mientras duerme o si alguien puede tener experiencias conscientes cromáticas, puede haber más o menos dificultades empíricas para responderla, pero parece claro que se puede establecer la diferencia entre una respuesta negativa y una respuesta positiva, sin que se plantee una cuestión filosóficamente interesante. Por otra parte, resulta significativo que Seager asimile el tipo de argumento que propone Dennett a una forma de desafío escéptico. Esta asimilación refleja en qué medida el compromiso con el realismo respecto a los estados mentales conscientes forma parte de la caracterización inicial de la experiencia consciente que se trata de poner en cuestión; de hecho, Seager no parece tener ningún problema en aceptar el argumento de Dennett como un argumento escéptico, dotado de un mayor o menor interés que cualquier otra hipótesis escéptica. Sin embargo, otra diferencia fundamental respecto a la argumentación de Dennett es que quien cumpliría con el rol correspondiente del escéptico no admite que se ponga en duda la existencia de la experiencia consciente, del modo en que se pone en duda la existencia del mundo externo, sino que trata de negar la inteligibilidad de la aplicación de un concepto en ciertas circunstancias.

Hasta cierto punto, la utilización del experimento phi por parte de Dennett como un modo de ilustrar su argumentación puede llevar a confusión en la medida en que no determina de un modo claro si su argumentación tiene un carácter empírico o lógico. Quizá el hecho de que este tipo de reflexión genere confusión todavía refleja una concepción de los problemas filosóficos y científicos muy compartimentada y separada. En este sentido, parece que el planteamiento de Chalmers se encuentra blindado ante cualquier hallazgo empírico, de manera que tomar en serio el problema de la conciencia ya implica, de algún modo, rechazar cualquier dato proveniente de la experimentación. Por otra parte, la argumentación de Dennett no tiene un alcance meramente empírico, sino que apunta a la falsedad de la concepción metafísica del Teatro Cartesiano, pero también señala el fracaso

o limitación de determinada vía de investigación. Si no hay un punto del cerebro al que se vincule el acaecimiento de la experiencia, cualquier programa de investigación que se comprometa con el materialismo cartesiano estará condenado al fracaso, del mismo modo que una profundización empírica en la física de los cuatro elementos nunca podría dar cuenta de todos los fenómenos que explica la física moderna. La idea de que no existe una evidencia bruta a la que apelar, completamente independiente de nuestras teorías y estructuras conceptuales, tiene ya un largo recorrido y, en términos generales, es difícilmente rechazable. La aceptación de esta idea, por otra parte, difumina los límites entre una investigación conceptual y una investigación empírica. A continuación, presentaré un ejemplo que muestra cómo la argumentación de Dennett, a pesar de basarse en información empírica, puede ampliar el espacio de lo que consideramos lógicamente posible. El argumento de Dennett no depende, exclusivamente, de cuál sea la interpretación correcta del fenómeno phi, sino que éste supone una ilustración de cómo la concepción estándar de la experiencia consciente no nos permite entender bien ciertos fenómenos.

5.4.4. La Ilusión de Capgras

Daniel Dennett presenta una reflexión, en *Sweet Dreams* (2005), respecto a la cuestión de si los *qualia* son los responsables de que la vida sea digna de ser vivida. En dicha reflexión, se sirve de un experimento mental que toma como punto de partida un extraño trastorno conocido como síndrome de Capgras. Las personas que padecen dicho trastorno, generalmente a consecuencia de una lesión cerebral, sufren la ilusión de que un familiar o un ser querido ha sido sustituido por un impostor; la creencia llega a ser tan vehemente que, en algunos casos, el sujeto que sufre del trastorno llega a agredir e incluso a asesinar al supuesto impostor. Según cierta interpretación del patrón ejemplificado en el síndrome de Capgras, este último es, en cierto sentido, la patología contraria de la prosopagnosia. Los pacientes de prosopagnosia muestran la incapacidad o dificultad para reconocer rostros, incluso de familiares, a pesar de que no tienen problemas para reconocer

objetos comunes. Una de las hipótesis que se barajan para dar cuenta de ambas patologías se vincula a la existencia de dos sistemas de reconocimiento en el sistema nervioso relativamente independientes: los sistemas cortical y límbico. El primero de ellos se encuentra vinculado a la conciencia, mientras que el segundo actúa a un nivel inconsciente. En el caso de la prosopagnosia, es el sistema cortical de reconocimiento el que se encuentra dañado; por el contrario, en el caso del síndrome de Capgras, es el sistema límbico el que no funciona adecuadamente, de manera que el sujeto, debido a una respuesta emocional defectuosa, en lugar de sufrir de una deficiencia perceptiva, genera la extravagante creencia, resistente al razonamiento y la evidencia, de que un impostor ha sustituido a un ser querido.

Una de las cosas que Dennett destaca de la evidencia proporcionada por estos casos es la falta de fiabilidad de los argumentos basados en la concebibilidad para extraer conclusiones respecto a la mente humana; en este sentido, ¿quién podría predecir que una lesión en una región determinada del cerebro puede producir una creencia como la ejemplificada en el síndrome de Capgras?

Sirviéndose de este trastorno, Dennett propone un experimento mental que quizá no sea posible, aunque no hay evidencia de que no lo sea. Probablemente, hay una gran cantidad de lesiones cerebrales posibles que de hecho no se han llegado a dar dado su perfil complejo, pero que darían lugar a trastornos difíciles de imaginar antes de disponer de estudios empíricos que los detecten.¹⁰³ En dicho experimento, el señor Clapgras ha sufrido una inversión de la dimensión afectiva, podemos decir que vinculada al sistema límbico, que antaño asociaba a los colores, de manera que sus distintas aversiones y preferencias respecto a los colores se encuentran invertidas. Podemos imaginar que la preferencia que asociaba por vestirse con ropa negra, ahora se asocia a la ropa fucsia, o que el tono grisáceo y opaco del cielo de un día otoñal, ahora le produce las sensaciones agradables que en el pasado le producía el cielo azul despejado de un día de primavera. Por otra parte, las habilidades cognitivas y verbales del señor Clapgras permanecen intactas en lo que se refiere a los colores, de manera que identifica adecuadamente el

¹⁰³ Pensemos, por ejemplo, en el trastorno conocido como anosognosia que, en una de sus formas, consiste en no ser capaz de percibir que la mitad del propio cuerpo se encuentra paralizada.

nombre de los mismos, y pasa sin ninguna dificultad los test cromáticos habituales, como cualquier persona normal.

El problema que se plantea es el siguiente: si aceptamos, por hipótesis, que los *qualia* hacen la vida digna de ser vivida, como a menudo se ha señalado, deberíamos decir que en el caso del señor Clapgras se ha producido una inversión de los *qualia*. Ahora bien, ¿no resultaría más plausible afirmar que los *qualia* cromáticos han permanecido estables mientras que la dimensión afectiva asociada a los colores ha variado? Dennett trata de mostrar, siguiendo una forma de argumentar semejante a la presentada en la sección anterior, que no hay un modo inteligible de distinguir entre las dos hipótesis. Podría pensarse que, dado que el señor Clapgras conserva todas sus habilidades cognitivas respecto a la identificación y reconocimiento de los colores, debería darse prioridad a la hipótesis de que los *qualia* han permanecido estables, ¿pero qué razón podría aducirse para llegar a esta conclusión? ¿Acaso no es compatible con la concepción de la conciencia que Chalmers baraja la idea de que las habilidades cognitivas y verbales respecto a los colores permanezcan estables, mientras que tanto la dimensión afectiva asociada a éstos como la fenomenología varíen a lo largo de un espectro de posibilidades? Si la argumentación de Dennett es correcta, estas posibilidades, así como las hipótesis a las que dan lugar, no son inteligibles, ya que no pueden diferenciarse ni siquiera conceptualmente, no sólo desde un punto de vista de la tercera persona, sino de acuerdo con el testimonio del sujeto que tiene la experiencia.

La cuestión es que Chalmers, ante este tipo de argumento, siempre puede llevar a cabo un movimiento que parece contrarrestarlo, reformulando el problema de la conciencia, o bien desplazándolo a un lugar en el que permanece intacto. Como hemos visto, dado cualquier mecanismo, propiedad conductual, funcional, física o, de un modo más radical, dado un mundo en todos los aspectos objetivamente observables idéntico al nuestro, ¿por qué cada uno de estos ítems o una réplica física de nuestro mundo tendrían que estar acompañados de

experiencia?¹⁰⁴ ¿No es posible, al menos conceptualmente, que un sujeto se encuentre encerrado en un cuerpo con unas determinadas disposiciones reactivas que no se corresponden en absoluto con su vida experiencial “interior”? Al fin y al cabo, ha habido descubrimientos asombrosos respecto a la actividad mental de personas que, aparentemente, se encontraban reducidas a una vida vegetal; dichas personas, de acuerdo con algunos experimentos en los que se han llegado a asociar determinados tipos de respuestas cerebrales a preguntas formuladas, serían conscientes, al menos respecto a determinado tipo de estímulos. Del mismo modo, aunque improbable fácticamente, ¿no podría haber un tipo de lesión que produjera un desajuste acusado entre los movimientos corporales y la propia experiencia, de manera que el cuerpo de las otras personas nos engañara totalmente respecto a su vida mental, quizá de un modo parecido a como nos equivocamos al interpretar los gestos de otras personas en un país extraño?

El tipo de argumento desarrollado en este capítulo, aun si constituye una manera de dejar de mirar en la dirección equivocada, puede dejar cierto tipo de insatisfacción filosófica, en la medida en que no ofrece una respuesta directa a la pregunta formulada en el párrafo anterior; de hecho, ni siquiera nos permite extraer una idea mínimamente sustantiva de en qué pueda consistir gozar de experiencia consciente. En el siguiente capítulo, seguiré una vía de reflexión alternativa que tratará de determinar en qué consiste el valor que asociamos a la experiencia consciente. La argumentación no se dirigirá a contrarrestar los argumentos de Chalmers, en la medida en que éstos pueden reformularse de una manera que intuitivamente se sobreponga a las objeciones planteadas. Más bien, trataré de aproximarme, precisamente, a lo que hace de la experiencia consciente una dimensión que marca una diferencia fundamental y que convierte a ésta en algo, en principio, misterioso. Aunque pudiera parecer que con este enfoque simplemente se plantea un problema adicional de la conciencia, de los muchos que compondrían una larga lista, mostraré que hay una continuidad relevante entre el *hard problem* de Chalmers y las cuestiones que plantearé, de manera que la incapacidad de la

¹⁰⁴ En Chalmers (2007) puede encontrarse la reiteración de esta pregunta como el núcleo básico a partir del cual plantear el *hard problem* de la conciencia.

concepción de la mente de Chalmers para dar cuenta del valor de la conciencia restará interés a su posición, tanto desde un punto de vista filosófico, como en relación a su propósito de cimentar una ciencia de la conciencia.

Hasta ahora he tratado de mostrar la ininteligibilidad de los supuestos metafísicos asumidos por Chalmers en su argumentación a favor de la irreductibilidad de la experiencia consciente. No obstante, no he presentado una historia alternativa que contrarreste la fuerza intuitiva de la concepción de Chalmers. En este sentido, mostrar la incoherencia de una teoría no garantiza *per se* la comprensión de un problema, sino que lo puede hacer todavía más enigmático.

¿Cuál es el tipo de relación que mantenemos con nuestros propios estados mentales conscientes? Negar que haya cierto tipo de asimetría entre la primera y la tercera persona es una tesis demasiado fuerte, que además no es una consecuencia necesaria del rechazo de la relación de familiaridad. En cualquier caso, no he dicho nada respecto a cuál pueda ser esta relación. Por otra parte, la negación de una concepción discreta y realista de la experiencia consciente daría paso a una concepción relacional y anti-realista de la misma, de manera que, por una parte, la pregunta sobre si un determinado estado consciente se ha dado no siempre tendrá una respuesta definida, y tampoco tendrá, teniendo en cuenta la argumentación del Capítulo 2, unas condiciones de individuación independientes de ciertas reacciones corporales y, en general, de ciertos criterios públicos, incluyendo las condiciones que hacen significativas las atribuciones de experiencia consciente a terceras personas.

También en este caso, las intuiciones en el sentido contrario no quedan completamente contrarrestadas. Quizá hablar del conjunto de las experiencias conscientes como un tipo de estados bien definido es un error al que no nos hemos sustraído, y en este capítulo la argumentación se ha centrado en casos de percepción visual. ¿Pero qué ocurre, por ejemplo, con el dolor? ¿Podemos decir que el dolor es una experiencia relacional que depende de un contexto y que de la que no podamos decir inteligiblemente si en ciertas ocasiones se da o no se da?

William Seager suscita de un modo contundente una intuición contra la argumentación que he desarrollado en este capítulo.¹⁰⁵ En el análisis del fenómeno phi, se plantea que para determinadas escalas de tiempo no tiene sentido plantear la pregunta por la instanciación de una determinada propiedad fenoménica; esta escala de tiempo ronda los 500 ms. Seager propone que imaginemos a Dios creando un mundo físicamente idéntico al nuestro, pero con una duración inferior a 500 ms. En tal caso, ¿diríamos que los seres creados en dicho mundo no gozarían de experiencia consciente? Resulta difícil, desde un punto de vista intuitivo, aceptar que así sea. Para poder dar una respuesta a esta pregunta, es necesario llevar a cabo un análisis alternativo de en qué consiste la experiencia consciente y qué rol juega en nuestras vidas. En el siguiente capítulo, esbozaré unas líneas generales que permitan un enfoque alternativo de la comprensión de la experiencia consciente.

En este capítulo, he analizado la concepción del Teatro Cartesiano criticada por Dennett, vinculándola a la conjunción de una metafísica discreta de la experiencia consciente con una posición fuertemente realista respecto a la misma, y he tratado de mostrar el conflicto de esta conjunción con el principio de que, en el caso de la experiencia consciente, ser y apariencia coinciden. Ante la objeción de que esta forma de argumentar se sustenta en la presuposición de la verdad de algún tipo de verificacionismo, argumenté en contra de dicha idea y traté de mostrar que la argumentación de Dennett tiene un alcance no limitado al fenómeno phi. Finalmente, consideré algunas de las insuficiencias de la argumentación presentada, en la medida en que no logra contrarrestar las intuiciones de Chalmers, al no presentar una imagen alternativa, cuyo esbozo se presenta en el siguiente capítulo.

¹⁰⁵ Para el siguiente argumento ver Seager (1999: 126).

6. Experiencia Consciente y Valor

6.1. Introducción

En el capítulo anterior desarrollé un argumento en contra de la inteligibilidad de que la experiencia consciente tenga un carácter metafísico discreto, así como en contra de una posición realista respecto a la misma. Al final del capítulo, sugerí que, si bien válido, el argumento tenía un cariz negativo y no gozaba de una imagen intuitiva que contrarrestara claramente las intuiciones que respaldan los argumentos de Chalmers. Dicho argumento apunta a una posición antirrealista y relacional respecto a la experiencia consciente, pero no da ninguna indicación positiva respecto a las condiciones generales bajo las cuales tiene sentido atribuir experiencia consciente. Una cuestión de la que hablé brevemente al mencionar la ilusión de Capgras, y que no recibe un tratamiento explícito en la obra de Chalmers, es el del valor que asociamos a la experiencia consciente. Parece claro que la vida merece la pena de ser vivida en la medida en que gozamos de experiencia consciente. Esto se manifiesta claramente en el trato que dispensamos a los animales en función de su grado de conciencia; pensemos, por ejemplo, en la diferencia con la que consideramos a un perro y a un insecto, respectivamente. Del mismo modo, un ser humano en un estado vegetativo, aun si su organismo mantiene las constantes vitales necesarias para subsistir, no goza de una vida de la que pueda decirse que es digna de ser vivida, y lo mismo sucede con una persona que vive con un sufrimiento crónico intenso. Ahora bien, parece que la vida de un zombi no merece la pena de ser vivida, de manera que la compleja red de relaciones causales ejemplificada en los zombis, al igual que en un ser humano normal, no es *per se* una fuente de valor. Esto conduce a la extraña conclusión de que seres semejantes a nosotros en todos los aspectos que podamos apreciar podrían no merecer el tipo de consideración moral que nosotros merecemos. En este sentido, ¿habría alguna diferencia entre la consideración que merece un zombi y la que merece una piedra?

En este capítulo, trataré de mostrar que la concepción de la experiencia consciente que defiende Chalmers no sólo no puede dar cuenta del valor de la

experiencia consciente, sino que es incompatible con la posibilidad misma de que la vida sea digna de ser vivida. La argumentación que presentaré pondrá en cuestión la idea de que la discusión en torno al problema de la conciencia pueda dilucidarse en un terreno abstracto, al margen de un contexto amplio en el que la experiencia consciente, lejos de ser un epifenómeno o una especie de elemento adicional, se individúa de una manera relacional, es decir, sólo dentro del complejo de interacciones en las que el individuo se encuentra integrado con su entorno.

La idea de que el valor de la vida se encuentra ligado a la experiencia consciente es difícilmente discutible, entendida en un sentido general. El siguiente texto de Antonio Damasio resume de un modo claro cuáles son los rasgos que hacen de la experiencia consciente una dimensión fundamental de la vida:

Sin conciencia—esto es, una mente dotada de subjetividad—no tendrías manera de saber que existes, por no hablar de saber quién eres y qué piensas. Si la subjetividad no hubiera empezado, aunque muy modestamente al principio, en criaturas vivientes mucho más simples que nosotros, probablemente la memoria y el razonamiento no se habrían expandido de la forma prodigiosa en que lo hicieron, y no se habría pavimentado el camino evolutivo al lenguaje y la elaborada versión humana de la conciencia que ahora poseemos. La creatividad no habría florecido. No habría habido ninguna canción, ni pintura, ni literatura. El amor jamás habría sido amor, sólo sexo. La amistad habría sido mera conveniencia cooperativa. El dolor nunca se habría convertido en sufrimiento—lo cual no es malo, piénsalo—aunque sería una ventaja dudosa puesto que el placer tampoco se convertiría en dicha. Si la subjetividad no hubiera hecho su aparición radical, no habría habido conocimiento ni nadie para darse cuenta de ello y consecuentemente no habría habido una historia de lo que hicieron las criaturas a través de los tiempos, ni cultura en absoluto. (Damasio 2010: 4)

El concepto de conciencia que en este texto se emplea, sin ser demasiado específico, se vincula a la noción de subjetividad y engloba tanto aspectos cognitivos como los aspectos que Chalmers consideraría fenoménicos. La consideración de la experiencia consciente como un elemento esencial en todas las dimensiones a las que se refiere Damasio parece integrar los dos aspectos de lo

mental que Chalmers distingue, de manera que el valor que atribuimos a la experiencia consciente no parece vincularse específicamente con el aspecto fenoménico de la experiencia.

La ruta que me propongo seguir a continuación trata, precisamente, de hacer justicia a la intuición general de que la experiencia consciente goza de un valor especial. Trataré un problema independiente y paralelo al “hard problem” que no podrá recibir una solución satisfactoria desde una posición como la de Chalmers. Al articular una posición que dé cuenta de esta dimensión de la experiencia, mostraré cómo la concepción de la experiencia consciente vinculada al dualismo de Chalmers, si bien puede siempre llevar a cabo un movimiento retórico que parece replantear el problema de la conciencia, no consigue recoger el aspecto que hace de la experiencia consciente algo importante y valioso, de manera que, independientemente de que Chalmers pueda responder a los argumentos presentados en los capítulos anteriores, sólo lograría una victoria pírrica, en la medida en que se comprometería con una concepción de la experiencia consciente con poca relevancia explicativa respecto a cuestiones fundamentales, generando problemas quizá más profundos que los que pretende conjurar.

En primer lugar, analizaré algunos casos de estados mentales inconscientes con el fin de determinar la especificidad de la experiencia consciente, y estudiaré un tipo de respuesta intuitiva a esta cuestión que, en sus rasgos generales denominaré detectivismo. Seguidamente, mostraré la insuficiencia de esta posición para dar cuenta del valor de la experiencia consciente y, en la misma medida, de la posición de Chalmers. En la sección 6.4, esbozaré un marco en el que puedan tratarse estas cuestiones de un modo más satisfactorio. Finalmente, analizaré la propuesta de ciencia de la conciencia de Chalmers respecto a su poder explicativo.

6.2. El Estudio de la mente y lo inconsciente

Uno de los argumentos más célebres contra el funcionalismo como una teoría de la mente que puede reducir o dar cuenta de todos sus aspectos es el célebre experimento mental del duplicado funcional de un cerebro humano implementado

en un sistema extravagante, como puede ser el conjunto de una nación.¹⁰⁶ En principio, aunque dicho sistema emule adecuadamente el conjunto de conexiones y redes causales que, por definición, constituiría una mente, tiende a imponerse la intuición de que dicho sistema no gozaría de experiencia consciente, en un sentido fenoménico, o al menos ese sería el propósito del argumento.¹⁰⁷ La cuestión es, ¿por qué tenemos la intuición de que dicho sistema no tiene, por ejemplo, *qualia* de dolor?

Al margen de que lo que muestra Dennett cuando habla de bombas de intuición defectuosas para referirse a algunos experimentos mentales es que no es evidente que en dichos experimentos se conciba un estado de cosas coherente, o que se imaginen adecuadamente todas las variables relevantes para poder evaluar el experimento, quizá no sea necesario recurrir a este tipo de escenarios para evaluar nuestras intuiciones: en la dimensión inconsciente de la mente podemos encontrar ejemplos reales de casos que constituyen ocasiones para evaluar nuestras intuiciones. En algunos casos, la complejidad de un comportamiento inconsciente puede llegar a ser aparentemente indistinguible de la de un comportamiento consciente, de modo que parece una estrategia prometedora evaluar nuestras intuiciones respecto a casos de comportamientos inconscientes.

Además, el estudio de la dimensión inconsciente de la mente puede resultar fructífero por otra razón. Un rasgo fundamental de la relación que mantenemos con nuestra propia mente es que el tipo de acceso que tenemos a la misma es especial, en la medida en que gozamos de una autoridad de la que carecen los demás. Parece claro que, en general, si queremos saber cuál es el deseo de una persona, o qué es lo que siente, el mejor método es preguntarle a dicha persona.¹⁰⁸ En cierto modo, Chalmers trata de atrapar este rasgo de lo mental al apuntar a la relación de familiaridad (*acquaintance*) como el tipo de relación que describe adecuadamente el vínculo que mantenemos con nuestra propia experiencia consciente. En el caso de los estados mentales inconscientes, este tipo de autoridad no se encuentra

¹⁰⁶ Para una formulación de este argumento ver Block (1978)

¹⁰⁷ Por supuesto, no se está sugiriendo que los individuos que componen el sistema carecen de experiencia consciente, sino el sistema constituido por éstos como tal.

¹⁰⁸ Ésta forma de caracterizar la autoridad de la primera persona puede encontrarse en Finkelstein (2003).

presente, de manera que el acceso del sujeto a dichos estados es similar al que otras personas puedan tener, recurriéndose en ambos casos, como veremos, a cierta evidencia.

Por otra parte, hablar en general de la dimensión inconsciente de la mente resulta sumamente ambiguo, y un tratamiento extenso de la misma excedería los propósitos de este trabajo. Podríamos distinguir muchos tipos de estados mentales inconscientes, que además lo son por causas muy diversas. Algunos de estos estados pueden asociarse a patologías neurológicas graves, y otros comparten, en ciertos aspectos, rasgos con lo patológico, aunque de un modo poco claro y comparable a los casos, por ejemplo, de lesiones cerebrales.; hay otros que, lejos de constituir una patología, forman parte de la vida mental cotidiana y habitual. Un ejemplo de este último caso son los momentos en los que no prestamos demasiada atención a nuestro entorno, al estar realizando algún tipo de actividad, a pesar de que lo percibimos de un modo fácilmente constatable. Pensemos en la actividad de conducir mientras mantenemos una conversación con nuestro acompañante, de modo que al final del trayecto no recordamos bien los detalles del mismo. Un ejemplo del primer caso es el trastorno conocido como visión ciega; las personas que sufren esta incapacidad, generada por una lesión cerebral, afirman no percibir nada en determinada región de su campo visual, a pesar de que sí parecen dar muestras de haber percibido adecuadamente el estímulo al plantearles ciertas cuestiones. Un tercer tipo de estados inconscientes se caracteriza porque, en principio, no se asocia a ningún tipo de lesión detectable, aunque comparte ciertos rasgos con estados patológicos en la medida en que se manifiestan como una anomalía respecto a lo que consideramos que es el modo normal de un sujeto para relacionarse con sus propios estados mentales; ejemplos de este tercer tipo son los célebres trastornos detallados por la bibliografía psicoanalítica, en los que el sujeto no tiene un acceso, debido a un proceso represor, a determinado estado mental, generalmente un deseo, aunque éste se manifiesta de un modo fragmentario y aparentemente inconexo en determinado tipo de conductas que dificultan o bloquean ciertas esferas de la vida del sujeto.

Esta enumeración no pretende ser exhaustiva, ni tampoco trata de dar una comprensión general de la dimensión inconsciente de la mente, sino que se trata de tres ejemplos escogidos con el fin de poder evaluar comparativamente algunos de los rasgos que hacen de un estado mental un estado consciente. A continuación, analizaré estos ejemplos extrayendo ciertas conclusiones respecto al tipo de relación que puede establecerse entre éstos y la teoría de Chalmers.

6.2.1. Visión ciega

El trastorno conocido como visión ciega¹⁰⁹ se asocia a una lesión en el córtex cerebral que afecta al procesamiento de la información visual, si bien dicha información puede procesarse en cierta medida. Aunque estos pacientes no ven nada, o afirman no ver nada, si por ejemplo se proyecta una luz de determinado color en la región afectada del campo visual, al proponérseles que elijan entre distintos estímulos que han podido percibir, eligen correctamente.

La visión ciega ha sido considerada en las últimas décadas como uno de los fenómenos que más problemas plantea a una concepción funcionalista de la experiencia consciente. De acuerdo con esta interpretación, en el caso del sujeto que sufre de visión ciega se produce un procesamiento de la información visual que no se encuentra acompañado de conciencia. En este sentido, John Searle plantea la siguiente pregunta:

[P]uesto que el paciente está recibiendo una información muy parecida tanto en la visión consciente como en los casos de visión ciega, ¿cuál es exactamente la diferencia, neurológicamente, entre la visión ciega y la visión consciente? ¿Qué añade el sistema cuando la misma información es aprehendida en una forma consciente? ¿Cómo entra la conciencia en la visión? (Searle 1997: 200)

Dada esta interpretación del trastorno, podría pensarse que la visión ciega es un caso real de un zombi, al menos respecto a una parte de la percepción visual. Dado que el rol funcional que juega el estado perceptivo correspondiente es idéntico tanto

¹⁰⁹ Para un tratamiento detallado del trastorno véase Weiskrantz (1986).

en un sujeto normal como en uno que sufre el trastorno, pero en el caso del paciente no hay conciencia, parece que la visión ciega supone un caso excelente para determinar cuál es la propiedad que marca la diferencia respecto a la conciencia. Por otra parte, si la visión ciega se interpreta como una objeción al funcionalismo, también supone una dificultad para la teoría de Chalmers. Inicialmente, podría pensarse que, por el contrario, dicho fenómeno fortalece la argumentación de Chalmers en la medida en que apunta a un aspecto de lo mental que no se deja atrapar en una red causal. Sin embargo, el tipo de dualismo propugnado por Chalmers es sensible a muchas de las objeciones que se le plantearían a un funcionalista duro. Si la visión ciega fuera una instancia de una determinada propiedad funcional que no se encuentra acompañada de la propiedad fenoménica correspondiente, entonces la supervinencia natural de la experiencia consciente sobre la base física no se mantendría, y no sería posible una ciencia empírica de la conciencia.¹¹⁰ Pero según Chalmers, esta interpretación de la visión ciega no es adecuada:

Concretamente, los sujetos con visión ciega parecen carecer del tipo de acceso usual a la información disponible. Su acceso es curiosamente indirecto, como lo atestigua el hecho de que no se encuentra fácilmente disponible para informar verbalmente, ni para el control deliberado de la conducta. La información se encuentra disponible para muchos menos procesos de control que en la información perceptiva estándar; se puede hacer disponible para otros procesos, pero sólo por medio de métodos poco usuales como la inducción o la elección forzada. De manera que esta información no cuenta como disponible directamente para el control global y los sujetos no son realmente conscientes de la información en el sentido relevante. La ausencia de experiencia se corresponde directamente con la ausencia de conciencia (*awareness*). Es también posible, quizás, que los sujetos con visión ciega tengan un tipo débil de experiencia, en cuyo caso podría querer decirse que tienen un tipo débil de conciencia, dándose apropiadamente los estándares de tener un carácter global y directo. (Chalmers 1996: 227)

¹¹⁰ Como veremos en el siguiente capítulo, si esta interpretación fuera correcta, la visión ciega constituiría un contraejemplo a los principios de coherencia entre la cognición y la fenomenología, los cuales constituyen los cimientos de la ciencia de la conciencia propuesta por Chalmers.

La respuesta de Chalmers es, en principio, adecuada, ya que la visión ciega no puede interpretarse como un caso real de ausencia de conciencia en el sentido psicológico o cognitivo, ya que no puede decirse que alguien que sufra de visión ciega ejemplifique las propiedades funcionales relevantes. Sin embargo, dicha respuesta no resulta efectiva en conjunción con algunas de las tesis que Chalmers defiende. Por una parte, Chalmers señala que los pacientes de visión ciega tienen un acceso indirecto a la información visual, que se aleja de la manera estándar de acceso a la información y que, además, se hace accesible a otros procesos a través de métodos inusuales. El problema es, ¿cómo puede Chalmers evaluar la relevancia, o irrelevancia, de todas estas variables? Si bien el acceso a la información visual es extraño, ¿por qué esto impide que haya experiencia fenoménica? O preguntando en la otra dirección, ¿por qué la manera usual y normal de acceder a la información visual da lugar a experiencia consciente? La cuestión no es que no dispongamos del conocimiento necesario para responder estas preguntas, sino que el marco que propone Chalmers para estudiar la conciencia excluye la posibilidad de responderlas. En última instancia, los hechos respecto a la experiencia fenoménica son irreductibles, y se vinculan con los hechos físicos de una manera bruta y contingente; esta tesis no sólo imposibilita saber cuándo un determinado conjunto de procesos físicos es un correlato de experiencia consciente, sino que impide, igualmente, comprender por qué la manera normal de relacionarse con los propios estados mentales se encuentra acompañada de fenomenología.

Por otra parte, Chalmers señala que carecemos de un acceso a los hechos relevantes para dirimir la cuestión, pero, ¿cuáles son estos hechos? Probablemente Chalmers se refiere a los hechos fenoménicos. En este caso, con independencia de que la crítica al Teatro Cartesiano funcione, Chalmers tiene que dotar de contenido al tipo de hechos irreductibles a los que apela, aunque esto es algo que no se puede hacer desde el punto de vista de la tercera persona, y para lo que el testimonio del propio sujeto no aporta nada, de manera que de poco servirá apelar a la relación de familiaridad con nuestros propios estados mentales conscientes. Este problema se puede plantear de un modo que se aproxima a una paradoja: para poder saber si

en los casos de visión ciega hay fenomenología, el sujeto tendría que dar un testimonio de su experiencia, pero entonces ya no sería un caso de visión ciega. Por lo tanto, el conocimiento de los hechos relevantes es incompatible con la existencia de los casos de visión ciega, y los casos de visión ciega, si de hecho se dan, excluyen la posibilidad de determinar los hechos relevantes.¹¹¹

En cualquier caso, aun reconociendo las limitaciones de la respuesta de Chalmers, es posible que la evidencia aportada por los casos de visión ciega sea demasiado fragmentaria como para poder establecer conclusiones claras. En este sentido, quizá sea más interesante que apelemos a un modo no patológico de relacionarnos con la dimensión no consciente de lo mental, con el fin de comprender lo que hace especial a la experiencia consciente en situaciones cotidianas.

6.2.2. Lo inconsciente y lo cotidiano

Frecuentemente, la dimensión inconsciente de lo mental se asocia a las investigaciones psicoanalíticas, o incluso a técnicas de control mental sobre aspectos de la personalidad que el sujeto no controla, de manera que quizá haya una imagen extendida de lo inconsciente como algo que irrumpe en la actividad normal del sujeto, alterándola o bloqueándola. Sin embargo, un aspecto que quizá esta imagen no tiene en cuenta demasiado a menudo es la importancia de que haya una parte inconsciente de la mente que no sólo existe como un fenómeno que de una manera habitual y cotidiana se encuentra presente, sino incluso como cumpliendo una función biológica. Muchas funciones básicas del organismo se encuentran, afortunadamente, fuera del control voluntario y consciente del individuo (pensemos en el latido del corazón, parcialmente la respiración o la digestión) aunque se encuentren vinculadas de múltiples formas a la actividad mental. Podría decirse que este tipo de procesos se encuentra en el lado opuesto a una línea

¹¹¹ Obsérvese que no se está presuponiendo la posibilidad de que la experiencia fenoménica no se corresponda con las propiedades funcionales relevantes, de acuerdo con leyes naturales, sino que aun dándose el tipo de necesidad natural con la que Chalmers se compromete, igualmente sería imposible dirimir la cuestión.

imaginaria que separara lo subpersonal de lo consciente. Sin embargo, no es tan evidente que pueda trazarse claramente dicha línea.

Parece que una posición realista respecto a los estados mentales conscientes sostendría que dicha línea existe en un sentido metafísico, aunque no seamos capaces de trazarla en algunos casos, como en el fenómeno phi. Al fin y al cabo, la tesis central de Chalmers es que los hechos referentes a la experiencia fenoménica han de añadirse al resto de hechos, y Dios vería claramente la línea divisoria que deja a un lado la fenomenología. Por el contrario, si negamos la tesis realista respecto a los *qualia*, tendría que ser posible la existencia de estados para los que no sólo no tengamos criterios sobre si son conscientes o no, sino para los que, igualmente, no tenga un sentido claro dicha pregunta. El caso del fenómeno phi tenía en cuenta determinadas magnitudes de tiempo, a partir de las cuales no resultaba inteligible una investigación que dirima entre teorías comprometidas con unos hechos determinados respecto a la conciencia. Si este resultado puede extrapolarse, en general, a la vida mental, de manera que no suponga un problema que surja exclusivamente a propósito de situaciones caracterizadas por incluir escalas de tiempo poco usuales, debería haber otros ejemplos de estados mentales de los que no pueda decirse si son conscientes o no.

En el listado de estados mentales que Chalmers presenta como perteneciendo, al menos parcialmente, a la dimensión consciente de lo mental, se encontraban elementos tan dispares como el olor a basura y la sensación producida al escuchar una fuga de Bach. Al tratar de determinar qué podían tener en común los estados que componían la lista, sugerí que no era sencillo determinar de un modo preciso una propiedad compartida por todos ellos. Según Chalmers, algunos estados gozaban de un componente fenoménico más claro que otros en los que predomina el aspecto psicológico o funcional. Un ejemplo del primer caso es una sensación olfativa, y un ejemplo del segundo es una creencia. En última instancia, en qué medida consideremos si uno de los estados es funcional o fenoménico constituirá una cuestión terminológica, y no revestirá mayor interés, siempre y cuando distingamos ambos componentes en las discusiones filosóficas.

Una característica notable de este listado es que parece que la clasificación de los distintos estados con aspecto fenoménico genera una categoría que difícilmente puede ser, a su vez, subdividida en categorías de un orden inferior que resulten relevantes en la formulación del problema de la conciencia. Por supuesto, los distintos estados mentales conscientes pueden ser clasificados en función de su origen, o teniendo en cuenta el tipo de estado mental que son (estados perceptivos, emociones, estados intencionales), pero estas subcategorías no parecen ser relevantes para la formulación del problema. Como señala Chalmers, “Al tratar los misterios filosóficos asociados con la experiencia consciente, una simple sensación de color plantea los problemas tan profundamente con la experiencia que se tiene con una coral de Bach” (Chalmers 1996: 11).

Sin embargo, hay numerosos ejemplos en los que no es evidente que un estado sea consciente o no, o que lo sea en la misma medida que en un caso paradigmático. Por supuesto, podríamos referirnos a los estados alterados de conciencia inducidos, por ejemplo, por drogas. También, las transiciones entre la vigilia y el sueño, o los mismos sueños, plantean dudas respecto a cómo deberían clasificarse; por ejemplo, ¿puede decirse, por ejemplo, que la experiencia asociada a la situación de yacer sobre la arena de una playa en un día soleado es idéntica en su fenomenología a la experiencia de soñar que yacemos en una playa en un día soleado?¹¹²

A continuación, me centraré en un tipo de experiencia que quizá constituya lo más parecido a una instancia real de un zombi, siendo, a la vez, una experiencia cotidiana. David Armstrong¹¹³ llama la atención sobre el fenómeno de la, por él denominada, conducción automática. Armstrong se refiere a la experiencia que puede tenerse al conducir un vehículo durante largo tiempo. En estos casos, a veces se tiene la sensación de que durante un gran período de tiempo no hemos sido conscientes de la conducción ni, hasta cierto punto, de lo que se ha ido percibiendo.

¹¹² Austin (1962) en *Sense and Sensibilia* llamó la atención sobre un tipo de falacia cometida al equiparar el carácter indistinguible de dos experiencias con su identidad cualitativa, de manera que, de acuerdo con dicha falacia, el hecho de que no podamos distinguir una experiencia onírica de una experiencia perceptiva implica que ambas experiencias son cualitativamente idénticas.

¹¹³ Véase Armstrong (1980).

Por supuesto, hay un sentido en el que se ha percibido adecuadamente el entorno y se han llevado a cabo una serie de acciones con cierto control voluntario (comparemos el caso de la conducción automática con el hipotético caso de un conductor sonámbulo). Sin embargo, diríamos que en estos casos hay un elemento que no se encuentra presente. El caso de la conducción automática es un ejemplo entre otros muchos, podemos pensar también en otros casos cotidianos como cuando alguien se dirige al frigorífico y al llegar al mismo no recuerda con qué fin se dirigía al mismo. Ahora bien, ¿cómo debemos entender estos casos, acaso es un modo de actividad funcional adecuada carente de experiencia fenoménica adicional? Y si respondemos afirmativamente a esta pregunta, ¿en qué consiste, precisamente, gozar, además, de experiencia consciente? La respuesta que Armstrong da a esta pregunta se compromete con una determinada imagen de la mente que aspira a respaldar, a su vez, un proyecto de investigación:

Mi propuesta es que la conciencia, en este sentido de la Palabra, no es nada, sino la *percepción o el darse cuenta del estado de nuestra mente*. El conductor en un estado de automatismo percibe, o se da cuenta de, la carretera. Si no lo hiciera así, el coche acabaría en la cuneta. Pero él no es en este momento consciente de esa conciencia de la carretera. Percibe la carretera, pero él no percibe su percibir, o cualquier otra cosa que está sucediendo en su mente. Él no es consciente, como lo somos normalmente, de lo que está sucediendo en su mente.

[...] El conductor en estado automático es tal que su “ojo interior” está cerrado: que no es en este momento consciente de lo que está sucediendo en su propia mente. (Armstrong 1980: 198)¹¹⁴

En el caso de la respuesta de Armstrong, parece que lo marca la diferencia entre simplemente percibir un objeto y ser consciente de que se percibe un objeto es,

¹¹⁴ La concepción de la mente supuesta en este fragmento parece comprometerse con una tesis muy cercana a la del Teatro Cartesiano, de manera que, aunque en un sentido metafórico, como las comillas indican, hay una especie de órgano situado en un lugar o punto, y que en determinado momento no se encuentra funcionando (está cerrado). En este capítulo no profundizaré en la argumentación que trata de refutar el materialismo cartesiano, sino que más bien trataré de evaluar la capacidad que tienen las teorías afines al mismo para dar cuenta del valor que atribuimos a la experiencia consciente.

precisamente, el acceso cognitivo a un determinado tipo de estado interno que en el caso del conductor se encuentra ausente. En este caso, además, parece que el hecho de que podamos llevar a cabo ciertas actividades de un modo automático es un requisito indispensable para que podamos realizar la acción exitosamente. En este sentido, es interesante tener en cuenta el modo en que diferentes personas realizan determinadas actividades, dependiendo del grado de conocimiento práctico del que disponen.¹¹⁵ En el caso de la conducción, el conductor novato o aprendiz necesita concentrar su atención en cada una de las actividades que lleva a cabo; por supuesto, este modo de proceder, aunque necesario como una etapa del aprendizaje, hace que la realización de la actividad por parte del aprendiz sea más torpe e ineficaz que en el caso del experto quien, por el contrario, gracias a la automatización de ciertas rutinas y procedimientos, es capaz de llevar a cabo los movimientos requeridos con mayor destreza y precisión y esto, a su vez, le permite descargar la atención de manera que la puede dirigir a la realización de otras actividades. En este contexto, parece que el hecho de que ciertas rutinas se lleven a cabo de un modo automático es una virtud, no sólo como una herramienta que facilita la supervivencia y maximiza la eficacia de nuestras acciones, sino como una forma de posibilitar nuevas formas de goce y placer. Por ejemplo, el placer asociado a la conducción se vincula a la dominación de los automatismos requeridos. Esto se ve de un modo más claro en el aprendizaje de un idioma: si se aprende un idioma en la edad adulta, es importante, al menos inicialmente, prestar atención a la misma estructura gramatical de la lengua y a aspectos de la misma que el usuario nativo o competente de la misma desconoce al resultarle transparentes. Sólo cuando se logra un alto grado de automatismo respecto al uso de las reglas de un idioma, es posible, por ejemplo, leer placenteramente la literatura escrita en el mismo.

Armstrong da cuenta de estos hechos apelando a la existencia de estados internos a los cuales podemos dirigir selectivamente nuestra atención, en caso de que así lo necesitemos, del mismo modo que podemos ser selectivos respecto a la percepción del mundo que nos circunda; en este sentido, el hecho de tener experiencia consciente se agota en la percepción de nuestros propios estados

¹¹⁵ Para un tratamiento de esta cuestión y de su relación con la experiencia consciente ver Nöe (2009).

mentales, mientras que la ausencia de experiencia consciente se explica en la medida en que no se tengan determinados estados, más allá de los estados perceptivos del mundo externo.

Quizá, en cierto sentido, este tipo de teoría, que como el mismo Armstrong reconoce tiene cierto carácter programático, difícilmente puede no ser verdadera. Pero esto no es debido a que se trate de una teoría autoevidente, o a que la evidencia empírica sea abrumadora al respecto. La idea de que se es consciente o no en virtud de encontrarse en cierto estado, más general que el tipo de teoría de la identidad propuesta por Armstrong aunque implicada por ésta, puede ser neutral respecto a múltiples interpretaciones del estatus de dicha teoría, y en este sentido es notablemente vaga. La expresión “en virtud de” es compatible con un sinnúmero de teorías metafísicas, de manera que sería realmente extraño que en casos en los que hay diferencias notables en cuanto al tipo de experiencias conscientes que tienen dos individuos, o incluso un mismo individuo, no se registre ningún tipo de diferencia cerebral o física. Pero este hecho, que es casi un supuesto de una ciencia sobre la mente, dista mucho de establecer una teoría definida sobre la misma.

Daniel Dennett también tiene en cuenta este tipo de casos cuando propone su propia teoría de la conciencia.¹¹⁶ Siguiendo la línea argumentativa en contra de la concepción del Teatro Cartesiano, la solución de Dennett se dirige a concebir un acercamiento a la experiencia consciente que no la reifique tratándola como un estado discreto; por otra parte, en el caso del conductor, resulta extraño considerar que éste no es consciente de lo que va percibiendo. Dennett apunta a que, precisamente, no hay un conjunto de hechos bien definidos respecto a la experiencia consciente que sean independientes conceptualmente de nuestros criterios de atribución de experiencia consciente: “no hay hechos fijos respecto al flujo de conciencia independientes de sondeos particulares” (Dennett 1991: 68). Si al conductor que realiza un largo trayecto se le pregunta por las cosas que percibe mientras conduce, puede dar una descripción detallada de las mismas y, seguramente recordará con posterioridad los detalles del trayecto. Por otra parte, parece que, si éste es el caso, tenemos una intuición robusta de que la experiencia

¹¹⁶ Véase Dennett (1991: 137-138).

fenoménica se encontraría presente. Un caso más claro que ilustra este tipo de explicación es la toma de conciencia del tictac de un reloj cuando éste cesa; si la explicación de Dennett es correcta, entonces no puede decirse que de un modo claro hemos sido conscientes del tictac del reloj hasta que el reloj deja de sonar, y si se nos preguntara por el ruido mientras éste se produce, igualmente repararíamos en él y seríamos consciente del mismo.

Por supuesto, desde un punto de vista como el de Armstrong, puede hacerse una lectura distinta de los hechos. Lo que ocurre cuando se pregunta sobre lo que se está percibiendo, o cuando cesa el tictac del reloj, es que la atención se dirige hacia los propios estados mentales, de manera que éstos se perciben, y en tal caso se puede hablar propiamente de experiencia consciente; además, estos estados se encuentran ahí y pueden ser percibidos o no. A continuación, presentaré un planteamiento alternativo respecto a las cuestiones discutidas que mostrará la insuficiencia de un planteamiento como el de Armstrong para dar cuenta de algunos de los rasgos relevantes de la experiencia consciente, así como de la concepción de la experiencia consciente de Chalmers.

6.3. Autoconocimiento y conciencia

Una cuestión de la que no he hablado todavía, y a la que Chalmers responde de una manera imprecisa, es la de la relación que mantenemos con nuestros propios estados mentales conscientes. Concretamente, planteé esta pregunta a propósito de la autoridad de la que gozamos en cuanto al conocimiento de nuestra mente. Como ya se señaló, parece que, si queremos saber algo sobre la vida mental de una persona, el mejor método para lograr este fin es preguntar a dicha persona. En un sentido amplio, esta autoridad es incuestionable, aunque, por supuesto, no puede ser aplicada en general a todos los aspectos de lo mental. Parece claro que no hay por qué tener una relación cognoscitiva respecto a las propias disposiciones psicológicas y que, en determinados casos, puede haber cierta opacidad respecto a estados intencionales como los deseos o las creencias. Sin embargo, ¿puede decirse lo mismo respecto a lo que Chalmers llama el aspecto fenoménico de lo

mental? Y si no es así, es decir, si no es posible que estemos equivocados respecto a nuestra propia fenomenología, ¿cómo puede darse cuenta de este hecho? Por otra parte, ¿en qué medida se vincula la fenomenología con el valor de la experiencia consciente? Como mostraré, las respuestas a estas cuestiones no son independientes entre sí.

El tipo de relación adecuada con nuestra propia fenomenología, según Chalmers, es de familiaridad (*acquaintance*); esta relación, definida de un modo impreciso, se puede considerar como una forma de acceso directo a la propia mente y constituye un tipo *sui generis* de relación cognoscitiva, en la medida en que no se caracteriza en términos de vínculos causales, y excluye la posibilidad de que respecto a la experiencia consciente se puedan introducir escenarios escépticos. Estos dos últimos aspectos se encuentran vinculados entre sí, ya que si la relación con nuestros estados mentales fuera causal y, por ende, contingente, sería posible construir argumentos escépticos respecto al conocimiento de éstos. Por otra parte, la manera de argumentar de Chalmers respecto a la irreductibilidad del aspecto fenoménico de la experiencia requiere también de cierta relación especial con la propia mente, en la medida en que sus argumentos reposan sobre un conocimiento de la experiencia que permite extraer, a priori, conclusiones respecto a lo que es metafísicamente posible.¹¹⁷

En la sección anterior, presenté una manera de entender la diferencia entre estados mentales conscientes y no conscientes vinculada a la existencia de dos tipos distintos de estados perceptivos. El caso de la conducción automática se explica por la ausencia de la percepción de determinados estados mentales, también perceptivos, pero vinculados a rasgos del mundo circundante, no del cerebro o, en general, del cuerpo; la percepción de estos últimos, a su vez, generaría otro tipo de estados que podrían identificarse con los estados mentales conscientes. Como ya he señalado, no pretendo argumentar en contra de la verdad, en cierto sentido, de este tipo de teoría, sino que más bien trato de comprender el estatuto de las tesis vinculadas a la misma y rechazo una interpretación última o

¹¹⁷ Como ya vimos en el capítulo tercero, el racionalismo modal de Chalmers, respaldado por el marco bidimensionalista que él mismo desarrolla, equipara la posibilidad lógica con la posibilidad metafísica.

metafísica de éstas. En mi argumentación, me serviré del estudio de un tercer tipo de estados mentales inconscientes, enumerados más arriba, que tienen en común con la visión ciega el hecho de que pueden considerarse patológicos, si bien no se pueden asociar a una lesión cerebral. Dichos estados tienen que ver, según la tradición psicoanalítica, con algún tipo de represión e irrumpen de un modo fragmentario y aparentemente arbitrario en el comportamiento de quienes los padecen. En la siguiente sección, mostraré en qué medida la teoría de Chalmers se encuentra limitada para enfrentarse a problemas suscitados por este tipo de estados mentales.

6.3.1. El detectivismo

Se tuvieron en cuenta, hasta ahora, dos ejemplos de lo inconsciente que tienen que ver, respectivamente, con un tipo de lesión cerebral y con la interacción normal entre nuestras acciones y la percepción. El caso de la visión ciega, como veíamos en la argumentación de Chalmers, no es un caso en el que se mantenga cierta organización funcional y no haya experiencia consciente. El caso de la conducción automática, según Armstrong, se explicaba por la existencia de una especie de mecanismo de detección interno, que puede estar activado o no. A continuación vamos a ver con mayor profundidad en qué medida ambos tipos de teoría pueden dar una explicación satisfactoria de la importancia de la experiencia consciente para nosotros.

En la tradición cartesiana, la autoridad respecto a los propios estados mentales goza de un vínculo conceptual con la misma noción de conciencia, manifestándose en la idea de que no podemos estar equivocados respecto a nuestros propios estados conscientes. Chalmers es capaz de dar sentido a la idea de que podemos equivocarnos respecto a nuestros propios estados conscientes, en la medida en que los dos conceptos de lo mental dan lugar a dos nociones de conciencia frecuentemente confundidas. Un zombi, al ejemplificar adecuadamente el mismo tipo de propiedades funcionales que un ser humano normal, es consciente en el sentido psicológico, aunque no lo sea, por hipótesis, en el sentido fenoménico.

Pero si nos ceñimos al sentido fenoménico de conciencia, no parece que tenga sentido distinguir una subclase de estados conscientes de los que, en cierto sentido, no somos conscientes; si nos parece que tenemos un dolor, entonces tenemos un dolor.¹¹⁸ El tipo de ambigüedad a la que apunta Chalmers, entre lo psicológico y lo fenoménico, no se aparta realmente de la tradición cartesiana, ya que el ámbito de lo fenoménico, aunque más restringido, conserva con pocas diferencias los rasgos centrales de lo que la tradición cartesiana considera como propiamente mental, en la medida en que los estados mentales conscientes de S no pueden ser desconocidos para S.¹¹⁹ Sin embargo, hay un rasgo de la concepción de la mente de Chalmers que igualmente se encuentra en su caracterización de la dimensión fenoménica, y que puede ser compartido por teorías muy dispares en otros aspectos, incluso en el caso de posiciones materialistas.

David Finkelstein, en su libro *Expression and the Inner* (2003), trata de desarrollar una teoría que dé cuenta de la autoridad de la que gozamos respecto a nuestros propios estados mentales, llevando a cabo un análisis crítico, a su vez, de los intentos más importantes de dar cuenta de dicha perplejidad. De entrada, el intento más atractivo de respuesta, y con mayor recorrido histórico, es lo que Finkelstein denomina una posición detectivista. Su caracterización general es la siguiente:

Un detectivista es alguien que cree que la habilidad de una persona para hablar de sus propios estados mentales tan fácil, precisa y autorizadamente como lo hace puede explicarse apelando al proceso por el que llega a saber de ellos. Según el detectivismo, puedo declarar mis propios pensamientos y sentimientos porque éstos son conscientes, y son conscientes gracias a un proceso cognitivo mediante el que he detectado su presencia. (Finkelstein 2003: 9)

¹¹⁸ Para la discusión de este capítulo dejaré de lado las dificultades a las que esta posición se tenía que enfrentar al combinar las tesis del Teatro Cartesiano con la relación de familiaridad.

¹¹⁹ Podría decirse que los zombis son esencialmente equivalentes a los animales no humanos, tal y como los concebía Descartes.

Apuntar a un vínculo conceptual entre experiencia consciente y algún tipo de cognición no supone un compromiso demasiado sustantivo con una teoría determinada, en la medida en que todavía no se ha dicho nada respecto al tipo de proceso cognitivo en particular con el que se relaciona la experiencia consciente, ni por qué un determinado proceso cognitivo, en lugar de otro, permite entender el hecho de que gocemos de autoridad respecto a nuestros propios estados mentales, dado que también accedemos a los estados mentales de los demás a través de un proceso cognitivo.

En este sentido, Finkelstein distingue dos formas básicas de detectivismo. La primera de ellas, denominada por él viejo detectivismo, se vincula a la tradición de la psicología de la introspección, entendida como un medio *sui generis* y exclusivo, para el sujeto, de relacionarse con los propios estados mentales. Finkelstein atribuye el ejemplo paradigmático de esta posición a Bertrand Russell en Russell (1912). El rasgo central de este detectivismo es que postula la existencia de un sentido interno (*inner sense*) que, en ciertos aspectos, puede ser entendido metafóricamente en relación con los sentidos que nos permiten conocer el mundo externo: “Tenemos familiaridad en la sensación con los datos de los sentidos externos, y en la introspección con los datos de lo que se puede llamar el sentido interno—pensamientos, sentimientos, deseos, etc.” (Russell 1912: 51).

La referencia a la relación de familiaridad (*acquaintance*) por parte de Russell, quien acuña dicho término en su sentido técnico, conecta directamente con la posición de Chalmers, si bien Chalmers utiliza el término con un significado más negativo, aunque parcialmente equivalente. Russell se compromete con una teoría de la percepción en la que los elementos perceptivos básicos son datos sensibles (*sense-data*), a partir de los cuales se construyen los objetos físicos mediante relaciones de inferencia, hablándose en tal caso de conocimiento por descripción. El conocimiento por familiaridad constituye, por otra parte, un conocimiento directo.¹²⁰

¹²⁰Chalmers no defiende explícitamente una teoría de los *sense-data*, e incluso algunas de sus tesis, como la paradoja del juicio fenoménico, podrían introducir una diferencia profunda entre el conocimiento de objetos físicos y el de estados fenoménicos, difícilmente asimilable a la concepción russelliana. Por otra parte, Chalmers se considera en algunos aspectos continuador de algunas de las tesis de Russell, principalmente en

Aun cuando el viejo detectivismo tome como referencia el modelo de la percepción, esta imagen debe interpretarse según ciertas restricciones. En primer lugar, en el caso del sentido interno, la relación cognoscitiva se agota en la familiaridad, de modo que todo objeto de este tipo es conocido mediante la misma y no se produce ningún tipo de inferencia, como en el caso de la percepción de lo externo. En segundo lugar, el conocimiento de nuestros estados mentales es un conocimiento cierto, para el que no cabe el error. Ésta es, quizá, otra manera de decir que el conocimiento introspectivo es directo, ya que no cabe en este caso una interpretación espacial de “directo”. El conocimiento del mundo externo es falible en la medida en que necesita de la inferencia, pero, ¿en qué podría consistir estar equivocado respecto a si tengo una experiencia de rojo o no?¹²¹

Este último rasgo convierte al mecanismo perceptivo implicado en la introspección en algo inmaterial, ya que ¿qué tipo de mecanismo perceptivo podría ser infalible y a la vez material? Cualquier órgano perceptivo puede funcionar inadecuadamente, tanto debido a un trastorno sistemático, como a una circunstancia concreta. En este sentido, ¿qué contenido debería dársele a la tesis detectivista más allá de una vaga referencia metafórica que no se sostiene en puntos cruciales? Por otra parte, la manera de construir la relación entre el conocimiento por familiaridad y el conocimiento por descripción aísla a la percepción de su contacto con el mundo, convirtiendo la existencia del mundo externo en algo problemático. No profundizaré, tampoco, en la dialéctica que este tipo de problemas genera. En cualquier caso, estas dos grandes dificultades suponen un problema especial para la concepción fisicista del mundo, en la medida en que remiten, por una parte, a una realidad inmaterial y, en el caso de la segunda, se cuestiona el estatuto de lo que deberían ser los hechos básicos del mundo.

Sin embargo, éstos no son problemas que una tesis detectivista tenga que afrontar. Esta posibilidad es la que da lugar a una formulación del detectivismo

Russell (1927), sobre todo en lo referente a la composición última de la materia. En este sentido, en una de las vertientes más especulativas de su obra, Chalmers propone una forma de materialismo en la que los componentes últimos de la materia serán un tipo de propiedades proto-fenoménicas. En cualquier caso, no profundizaré en la comparación detallada entre ambas teorías.

¹²¹ Esta pregunta no es retórica, ya que la idea de que no podemos estar equivocados respecto a nuestras propias experiencias no está exenta de objeciones.

exenta, en principio, de las dificultades del viejo detectivismo. Según ésta, la conciencia que tenemos de nuestros propios estados mentales se encuentra vinculada, igualmente, a un mecanismo de detección, a un proceso cognitivo, aunque en este caso el acceso a los estados mentales no será *sui generis*, de modo que no hay una diferencia esencial entre la percepción de objetos físicos y la de estados mentales. La satisfacción de este principio general es compatible con diversas teorías, dependiendo de la concepción que se tenga de los estados mentales. Finkelstein señala de nuevo a Bertrand Russell como un representante del nuevo detectivismo en Russell (1921). En dicha obra, Russell propone una teoría conductista del deseo, según la cual un deseo es un estado de inquietud, y aquello que pone fin a este estado es el objeto deseado. Una consecuencia de esta teoría es que tener un deseo no equivale a conocer el objeto del mismo, sino que lo llegamos a conocer a partir de experiencias pasadas, sirviéndonos en casos actuales, por tanto, de inferencias.¹²² Al margen de la implausibilidad de esta teoría, el rasgo que la convierte en una forma de nuevo detectivismo es que el modo que tenemos de conocer nuestros propios estados mentales no es *sui generis* respecto al modo en que conocemos hechos y sucesos externos. De esta manera, las dos objeciones que se planteaban respecto al viejo detectivismo dejan de ser aplicables. Sin embargo, el viejo detectivismo no paga un alto precio a cambio de nada; es precisamente la idea de que gozamos de una autoridad especial respecto a nuestros propios estados mentales lo que trata de reflejar la idea de la introspección, pero en el caso del nuevo detectivismo se niega que gocemos de una autoridad especial en ese sentido, de manera que niega la perplejidad que suscita esta reflexión.¹²³

Este modelo de nuevo detectivismo es en cierto sentido atípica, y no goza de defensores actuales. El requisito de comprender el acceso a los propios estados mentales ha recibido una nueva formulación en las últimas décadas, en relación a los avances en ciencia cognitiva y en el conocimiento del cerebro. Volviendo a la interpretación que Armstrong lleva a cabo del fenómeno de la conducción

¹²² También podemos encontrar una teoría de este tipo en Carnap (1932) y en Ryle (1949).

¹²³ Aunque en cierto sentido esta forma de detectivismo niega la autoridad especial de la primera persona, podemos encontrar algunos intentos de dar cuenta de la misma desde una posición conductista.

automática, lo que explicaba la diferencia entre estados conscientes y estados inconscientes, aunque acompañados de una percepción funcionalmente precisa, es que en el caso de los estados conscientes no sólo hay una percepción del mundo externo, sino que además hay una percepción adicional de los estados que se generan en el cerebro, y esta percepción no sólo no tiene nada que ver con un órgano inmaterial, sino que debe entenderse en el seno de un proyecto naturalista y, por lo tanto, debería ser posible desentrañar empíricamente el mecanismo de la misma.

Los detalles de esta variante del nuevo detectivismo no pueden ser completados a priori, ya que dependen de descubrimientos empíricos, pero el carácter naturalista de esta posición excluye la posibilidad de que seamos infalibles respecto a nuestros propios estados mentales. Esta posición no es sensible, por tanto, a la objeción de inmaterialidad o que se le puede hacer al viejo detectivismo. Respecto a cómo podría darse cuenta, en este marco, de la autoridad de la primera persona, no hay una respuesta que surja automáticamente. Finkelstein considera que el nuevo detectivismo no puede dar cuenta de esta autoridad, sin embargo, el punto de partida del proyecto excluye una solución a priori de esta cuestión, y más bien la desplaza al terreno de la ciencia, por lo que esta objeción, así planteada, no resulta demasiado importante. En cualquier caso, el nuevo detectivista puede apelar a un órgano perceptivo adicional, aunque no esencialmente distinto de otros órganos perceptivos, para explicar dónde reside esta autoridad. El propio sujeto es la mejor persona a la que preguntar sobre sus propios estados mentales, en principio, precisamente porque dispone de un mecanismo físico u órgano que los detecta. El contraste respecto al conocimiento de la mente de los demás podría entenderse a partir del contraste que hay entre percibir colores o tener un conocimiento de estos a través del testimonio de otros debido a una ceguera; en ambos casos se carecería de un órgano perceptivo adecuado.

Sin embargo, Finkelstein plantea una objeción que se dirige directamente al rasgo central de toda posición detectivista, sin distinguir entre la versión vieja o la

nueva.¹²⁴ Dicha objeción se presenta en forma de un diálogo imaginario entre un detectivista y un antidetektivista. A continuación, desarrollaré una interpretación de dicha objeción y extraeré algunas conclusiones relevantes para la discusión.

6.3.2. “Consciente de” y “conscientemente”

Como ya se señaló, parece que el problema de la conciencia puede reformularse de nuevas formas en función de los diversos intentos de respuesta propuestos. Cualquier intento de dar cuenta de la experiencia consciente en términos de entidades que no tengan que ver con estados mentales conscientes siempre tendrá que responder a la siguiente pregunta: ¿por qué un estado x de cosas determinado se encuentra acompañado de experiencia consciente?

En este capítulo he tratado de analizar en qué consiste que un estado sea consciente, a partir de ejemplos de estados inconscientes. He presentado un conjunto de posiciones que tratan de dar cuenta de en qué consiste que un estado sea consciente, equiparando la conciencia con la existencia de un mecanismo de detección, y a continuación desarrollaré una objeción dirigida directamente a este último rasgo, común a toda posición detectivista.

Hasta ahora, no se ha tenido en cuenta el tipo de estados inconscientes que tienen que ver con un trastorno que no se define, al menos de momento, en términos neurológicos. Se trata de aquellos estados mentales que han llegado a ser inconscientes en virtud de un proceso complejo asociado a la represión y a la fragmentación del yo, y que han sido objeto de un análisis profundo en la tradición psicoanalítica. Más allá de las controversias que suscite el estatuto de la ciencia psicoanalítica, parece difícil negar la existencia de estados mentales inconscientes, cuya inconsciencia responda a un conflicto personal que genera una represión, igualmente, parece difícil negar que estos estados inconscientes, no obstante su

¹²⁴ No es explícito que Finkelstein afirme que la objeción que sigue a continuación aplique para los dos tipos de detectivismo. El viejo detectivismo es descartado por su carácter inmaterial y anticientífico. Sin embargo, como mostraré a continuación, la objeción se dirige al rasgo general de toda posición detectivista.

condición, continúen manifestándose en la vida de la persona¹²⁵ y, en ocasiones, dichas manifestaciones irrumpen de un modo imprevisible e inconexo, bloqueando o dificultando la vida y el bienestar de la persona. Sin entrar en demasiados detalles, no parece afirmarse nada demasiado problemático al reconocer la existencia de casos en los que alguien no es consciente de que está enfadado con una persona, o casos más dramáticos en los que alguien profesa un amor inconsciente e inconfesable por determinada persona, y que dadas ciertas circunstancias nunca será correspondido.

Lo que caracteriza a una posición detectivista es que somos conscientes de nuestros propios estados mentales gracias a un mecanismo de detección, que nos permite tener un conocimiento de éstos. A diferencia del viejo detectivismo, el nuevo reconoce la posibilidad de que dicho mecanismo de detección yerre o no detecte algunos estados mentales, como en el caso de la conducción automática. Aceptando que ésta es la posición del nuevo detectivista, Finkelstein plantea la objeción que un antidetectivista podría plantear:

AD: Bueno, entonces, el primer paso para ver lo que está mal con tu forma de dar cuenta de la autoridad de la primera persona es trazar la distinción entre dos cuestiones: si un estado mental es conocido o no por su sujeto y si es consciente o no. Podría saber que estoy enfadado, incluso aunque mi enfado sea inconsciente. Imagina que mi terapeuta, en quien yo confío y tengo buenas razones para confiar, me dice que he estado inconscientemente celoso de mi hermana desde que era un niño. Podría, de este modo, aprender un hecho sobre mí mismo—que albergo celos inconscientes—sin llegar a ser conscientemente celoso.

Una manera de expresar esta idea sería decir que hay una diferencia entre “consciente *de*” y “conscientemente.” Puedo ser consciente *de*, i. e. conocedor de, mis celos sin ser conscientemente celoso—i. e. ya que para que mis celos sean conscientes y no simplemente algo *de* lo que soy consciente—no es suficiente para mí *saber* que estoy celoso. (Finkelstein 2003: 21)

¹²⁵ Por supuesto, un estado mental, aunque sea inconsciente, tiene que asociarse algún tipo de manifestación detectable, de otro modo ¿qué contenido podría asociarse a la idea de dicho estado mental existe?

La distinción entre “consciente de” y “conscientemente” apunta a que la autoridad de la primera persona, al menos en un sentido relevante, no puede explicarse sólo en términos de lo que se conoce. Ésta no es una limitación del detectivismo en una de sus formas concretas, sino que cuestiona, precisamente, que la detección sea el tipo de relación adecuado para captar la autoridad que tenemos respecto a nuestros estados mentales. A lo sumo, todo lo que podría mostrar es que nos relacionamos con algunos de nuestros estados del mismo modo que lo hacemos con los estados de los demás, y éste es el caso cuando somos meramente “conscientes de” nuestro enfado.

Para saber sobre los estados mentales de los demás necesitamos recurrir a cierta evidencia, al testimonio de la misma persona o de otros, a cierta conducta o reacción, etc. Pero si se nos preguntara a nosotros mismos cómo sabemos que tenemos sed, parecería como si la persona que lleva a cabo dicha pregunta no hubiera comprendido bien el uso de los enunciados sobre los estados mentales del propio hablante. Sin embargo, en el caso del enfado descubierto a través del diagnóstico del terapeuta, el tipo de relación que mantendríamos con dicho estado es, en lo relevante, similar a la que mantenemos con estados mentales de otras personas. Por una parte, éste es un caso especial que no puede servir de modelo extrapolable a la comprensión general de la autoridad de la primera persona; por otra parte, la distinción introducida por Finkelstein pone de relieve que la tesis detectivista no es lo suficientemente potente para poder explicar la diferencia entre las dos formas de conciencia.

El diálogo introducido por Finkelstein continúa con algunos intentos del detectivista de defensa de la relevancia de su tesis. En primer lugar, replica que en el caso de la persona que conoce su enfado, no lo hace de la manera apropiada, a través de una especie de “ojo interior”; pero, ¿por qué esta diferencia tendría que ser relevante? En cierta manera, la respuesta sólo pospone la comprensión del problema. El antidetectivista señala que la importancia de la conciencia no radica en aprender algo, ya que, volviendo al caso del paciente que visita un terapeuta, ¿qué sentido tendría seguir yendo al terapeuta para lograr llegar a estar “conscientemente” enfadado? Finkelstein señala que es cuando tenemos esta

relación adverbial con nuestros estados mentales cuando podemos decir que nuestras experiencias conscientes son valiosas, de manera que la vida de un hombre que se relacionara con sus estados mentales del mismo modo que se relaciona con los estados mentales de los demás carecería de algo importante y valioso, y tendría una vida superficial y empobrecida. El siguiente movimiento del detectivista es apelar a la fenomenología, y a la idea de que cada forma de percepción proporciona una fenomenología diferente, pero, una vez más ¿por qué la fenomenología goza del poder que hace a la conciencia especial?

El último intento del detectivista es apelar a la riqueza de aspectos que proporcionan los distintos modos de percepción como aquello que convierte a la conciencia en algo importante, y en el caso de que no detectemos nuestros estados mentales a través del sentido interno, entonces esta riqueza perceptiva estaría ausente. Sin embargo, ¿no podríamos imaginar una manera de tener acceso a toda esa riqueza aspectual a través de un mecanismo que no fuera el sentido interno, quizá observando las propias expresiones faciales o a través de un escáner detallado del cerebro, de manera que para poder conocer las propias emociones se necesitara estudiar el propio rostro en el espejo? El problema del detectivista no es que su teoría sea falsa, sino que ésta no dispone de elementos que permitan profundizar en la comprensión de la autoridad de la primera persona y ni la relación como la de familiaridad o ni un mecanismo supernatural de detección gozarían del poder explicativo que nos permitiera entender el valor que atribuimos a nuestras experiencias conscientes. En la siguiente sección, mostraré cómo se relaciona esta discusión con la teoría de Chalmers y esbozaré una manera alternativa de enfocar la cuestión.

6.4. Experiencia, pasividad y acción

Al presentar el detectivismo, establecimos algunos vínculos entre éste y la concepción de la mente de Chalmers. Sin embargo, no es tan claro cuál es el aspecto de dicha concepción, en virtud del cual la objeción de Finkelstein afecta también a ésta. La relación de familiaridad respecto a la propia mente con la que se

compromete Chalmers sitúa su posición más cerca del viejo detectivismo, aunque no está tan claro en qué medida puede aplicarse la idea de detección a una concepción tan negativa de familiaridad, que además es compatible con la paradoja del juicio fenoménico. Sin embargo, hay un rasgo todavía más general del detectivismo que lo hace sensible a la objeción presentada. El detectivista del diálogo de Finkelstein lleva a cabo un movimiento dialéctico recurrente, consistente en mencionar un ítem que por sí mismo debería tener el poder de dar cuenta de la autoridad de la primera persona y de aquello que es importante y confiere valor a nuestras vidas en relación con la conciencia. La cuestión es, ¿cómo la solución a estos problemas puede residir en una respuesta que simplemente señale a un elemento discreto como la clave a tener en cuenta?

Para Chalmers, la existencia de experiencia consciente se reduce a que se instancien ciertas propiedades con un carácter especial. Ahora bien, ¿qué significa que una propiedad fenoménica se instancia? ¿Podemos entender la ejemplificación de propiedades fenoménicas cromáticas a la luz de cómo el color rojo se instancia en un tomate? Esta última pregunta genera una ambigüedad difícil de desentrañar, en la medida en que las propiedades fenoménicas de las que habla Chalmers se instancian en estados mentales, y no en objetos.¹²⁶ En este sentido, ¿cómo podemos entender que una de las propiedades que se ejemplifica en un estado mental es una experiencia de rojo?

Como ya señalamos, el modelo general que se suele tomar para pensar en la fenomenología es el visual. Uno de los argumentos del detectivista del diálogo apelaba a la riqueza de aspectos que proporciona cada medio perceptivo; en este sentido, una persona que sufriera de una visión ciega en todo su campo visual carecería, supuestamente, de la riqueza de aspectos que confiere la percepción visual, aun si funcionalmente fuera indistinguible de una persona con una visión normal. Ahora bien, ¿es este tipo de visión ciega, no ya empíricamente posible, sino

¹²⁶ Stroud (2000) cuestiona que esta sea la única forma de entender la lógica de nuestras percepciones de color. Los colores pueden describirse como sensaciones, pero normalmente percibimos que un objeto tiene un color, no simplemente tenemos una sensación de color. Esto es lo que Stroud llama “ver proposicional”, y dada esta distinción, es bastante difícil que podamos dar cuenta de nuestras percepciones de color simplemente haciendo referencia

conceptualmente coherente? A continuación, trataré de mostrar que no y esbozaré un modo alternativo de pensar sobre la experiencia consciente.

Alva Nöe, en *Action in Perception* (2004), identifica una manera de concebir la experiencia visual con un gran poder intuitivo y que tiene implicaciones relevantes para esta discusión:

Cuando tratamos de entender la naturaleza de la percepción sensorial, tendemos a pensar en términos de visión, y cuando pensamos en la visión, tendemos a pensar que el ojo es como una cámara y que la visión es un proceso cuasi fotográfico. Ver, suponemos, es tener experiencias de tipo instantánea de la escena que tenemos delante de nosotros. Abres tus ojos y se te dan experiencias que representan la escena—como las de una imagen—con un enfoque nítido y un detalle uniforme del centro a la periferia. (Nöe 2004: 35)

Podemos llamar a la concepción de la experiencia que Nöe describe en este texto la *concepción fotográfica* de la experiencia. En el símil que describe Nöe, en una experiencia visual se presenta una serie de propiedades de un modo detallado y simultáneo y, del mismo modo que en el caso de una fotografía, parece que podemos hablar de un objeto bien definido con unas propiedades que están ahí, en cualquier caso. Esta concepción ilustra de un modo más claro la concepción realista de la experiencia de la que hablé en el capítulo anterior. Nuestras experiencias conscientes son como objetos con unas propiedades bien definidas que, no obstante su carácter efímero, en caso de que pudieran, de algún modo, congelarse, podríamos observarlos en detalle y que, precisamente, por ser efímeros, es posible que tengan propiedades que se nos escapan en el breve tiempo que tenemos para poder observarlos. Además, los objetos de estudio de una hipotética ciencia de la conciencia, de ser correcta la concepción fotográfica, podrían estudiarse del mismo modo que objetos físicos o partículas. En última instancia, la concepción fotográfica de la experiencia es una ilustración de la reificación de la experiencia consciente.

La concepción fotográfica de la experiencia debería ser compatible con una serie de hechos bien documentados empíricamente, especialmente en relación a la fisiología de la percepción. Por ejemplo, sabemos que, dada la fisiología del ojo, la

imagen retinal que posteriormente estimula el nervio óptico se encuentra invertida. Un ejemplo más intuitivo, en la misma línea, es que no disponemos de una única imagen en la retina, ya que tenemos dos ojos, pero experimentamos un único campo visual. Estos ejemplos podrían mostrar que nuestra percepción es una alucinación sistemática. Ahora bien, sugerir que la experiencia consciente pueda ser una ilusión parece contradictorio; siguiendo la tradición cartesiana, una persona puede tener la ilusión de que se encuentra en una playa, cuando realmente yace sobre su cama y sueña que está en una playa. Pero, ¿es posible que a alguien le parezca tener la experiencia de estar en una playa y que de hecho no la tenga? Quizá, entender bien esta pregunta requiere una reflexión más cuidadosa respecto a lo que significa “tener una experiencia”.

Un fenómeno sorprendente de nuestra percepción es que, dados ciertos hechos vinculados a la fisiología de nuestros órganos perceptivos, no nos percatamos de ciertas discontinuidades. Hay un hecho referente a la visión que socava la idea de que la percepción visual, en contraste con el tacto, tiene un carácter pasivo, y es que sólo se puede ver si el ojo se encuentra en movimiento, cuando efectúa los llamados movimientos sacádicos (los cuales se llevan a cabo una media de cinco veces por segundo); sin embargo, no percibimos ninguna discontinuidad en el intervalo entre dichos movimientos. Otro ejemplo conocido es que, debido a la inserción del nervio óptico, hay una parte de la retina que no dispone de células fotorreceptoras, de manera que no somos capaces de ver cierta región del campo visual. Sin embargo, no percibimos un hueco en nuestro campo visual, a modo de un agujero negro, sino que aparentemente nada perturba la continuidad de nuestro campo visual. Una explicación de este fenómeno es que el cerebro, en cierto modo, elabora la información que le proporcionan los órganos sensoriales. Este tipo de explicación parece presuponer, en cierta medida, la concepción del Teatro Cartesiano, e incluso se aproxima a cierta interpretación estalinista de la misma, en la que el cerebro actúa como una especie de editor que realiza montajes para ser “proyectados” posteriormente.

Sin embargo, si estos montajes se dieran, seríamos capaces de describirlos en detalle; pero esto, no sucede realmente. En Dennett (1991) se presenta un

ejemplo que ilustra este hecho. Si entramos en una habitación cuyas paredes se encuentran cubiertas por un papel en el que aparecen cientos de caras de Marilyn Monroe, al estilo Andy Warhol, en una fracción de segundo percibiríamos que, de hecho, hay cientos de copias de la foto de Marilyn; pero en dicho intervalo de tiempo, quizá sólo llegaríamos a tener una información visual adecuada de dos de las fotos. Según la idea de que el cerebro es un editor, podríamos pensar que el cerebro escanea las copias que propiamente vemos, y las reproduce un centenar de veces. La realidad, no obstante, es que si tuviéramos que dar una descripción detallada de una de las fotos que se encuentra alejada del ángulo que incide en el lugar en el que se encuentra la concentración apropiada de células fotorreceptoras, apenas podríamos describirla. Y tampoco daríamos una descripción detallada de un supuesto montaje estalinista. Pero entonces, ¿habría que decir que la experiencia consciente es una ilusión, en la medida en que percibimos mucho menos de lo que creemos percibir? Esta es una cuestión que requeriría un tratamiento más amplio del que se le puede dar en este capítulo; en lo que resta, me limitaré a señalar una posible fuente de confusión que oscurece el debate.

Para que se produzca un conflicto entre lo que creemos que percibimos y lo que de hecho percibimos, en la medida en que creeríamos experimentar una riqueza de aspectos cualitativos que de hecho no experimentamos, tendríamos que comprometernos con la idea de que aquello de lo que decimos que es rico en aspectos es nuestra experiencia. En cierto sentido, es difícil negar que así sea, pero una vez más el problema se encuentra en el estatuto que atribuimos a esta tesis. Teniendo en cuenta el tipo de situaciones reales en las que nos encontramos cotidianamente, es poco habitual que tengamos que describir con todo detalle una experiencia que sólo dura un fragmento de segundo; normalmente, más que describir nuestra experiencia, y especialmente en el caso de la percepción visual, lo que hacemos es describir el mundo, de manera que la riqueza de la fenomenología no es sino la riqueza de los distintos aspectos que podemos apreciar en nuestro entorno. Cuando decimos que nuestra experiencia visual es muy rica, estamos diciendo, en parte, algo sobre el mundo que nos circunda, y no simplemente sobre un estado interno de nuestro cerebro, o una especie de fotografía fija fenoménica

que podemos identificar con nuestra experiencia. Es precisamente esta confusión a la que apunta Alva Nöe en el siguiente texto:

La transparencia de la experiencia no es un obstáculo para la fenomenología concebida apropiadamente. Pero muestra que comprendemos mal la experiencia si pensamos en ella como ocupándose de la estructura del campo visual.

Experimentamos el mundo como siendo ilimitado y densamente detallado porque no habitamos un dominio de fijaciones visuales de tipo instantánea. Cuando mantenemos nuestra mirada fija de esa manera, no miramos alrededor, y en la medida en que no miramos alrededor, no vemos. La visión es activa; es una exploración activa del mundo. (Nöe 2004: 72)

En un mundo en el que sólo experimentáramos un torrente de sensaciones efímeras e inconexas sería muy difícil tener una concepción reificada de la experiencia consciente (al margen de que es difícil incluso imaginar que pudiéramos tener una concepción del mundo si todo fueran sensaciones inconexas). Si esta forma de pensar en la fenomenología es correcta, nuestra experiencia no es independiente conceptualmente de nuestro cuerpo y de la relación que éste mantiene con su entorno, de manera que la conciencia tiene un componente esencialmente activo, en contraposición a la concepción del Teatro Cartesiano.¹²⁷ Otra manera de expresar esta idea es que tenemos las experiencias conscientes que tenemos en virtud, tanto de las posibilidades que tenemos de interactuar con nuestro entorno, como del conocimiento adquirido en el pasado respecto a esas posibilidades.

Esta línea de argumentación también puede ayudar a entender la especificidad de nuestras distintas experiencias. El *hard problem* de la conciencia que plantea Chalmers no sólo se limita a la cuestión de por qué determinados procesos subpersonales están acompañados de conciencia, sino que aun si quedara garantizada la existencia de conciencia en algunos casos, todavía podría plantearse la pregunta por el carácter cualitativo de una experiencia, de manera que

¹²⁷ Precisamente, la idea del Teatro Cartesiano explota una intuición generada por una concepción pasiva de cómo funciona la percepción visual; como hemos visto, esta concepción se encuentra actualmente obsoleta, en la medida en que en la percepción visual hay un componente activo. Ahora bien, el tipo de vínculo entre nuestro y el entorno del que aquí se habla no es meramente causal

el hecho de que tengamos experiencias con determinado carácter cualitativo, y no otro, parece arbitrario. Esta arbitrariedad nos impediría saber, en parte, cómo es, por ejemplo, la experiencia consciente de un murciélago. Sin embargo, como ya señalamos con anterioridad, el ejemplo de Thomas Nagel delata, en cierta medida, que el tipo de experiencias de las que gozamos no es tan arbitrario. Si pensar en la experiencia de un murciélago para ilustrar el problema de la conciencia nos resulta convincente, es precisamente porque dado lo que sabemos de los murciélagos desde el punto de vista de la tercera persona, su forma de experimentar el mundo tiene que ser muy distinta a la nuestra. El hecho de que los murciélagos perciban el mundo, principalmente, a través de sonidos parece que tiene que vincularse con una fenomenología radicalmente distinta.

Esta idea nos lleva de nuevo a la réplica del detectivista, según la cual la persona que llega a conocer sus estados mentales observando sus expresiones faciales en el espejo no tiene el tipo adecuado de conciencia en la medida en que no se sirve de su “ojo interno”. Esta respuesta exigía una explicación de por qué determinada forma de percepción es relevante respecto a la distinción entre ser meramente *consciente de* algo y serlo *conscientemente*; teniendo en cuenta lo dicho, nos encontramos en situación de poder esbozar, si bien no una respuesta definida, al menos un marco en el que dicha distinción puede ser comprendida mejor. El individuo que se relaciona con sus estados mentales de un modo anómalo ejemplifica una red de relaciones entre sus estados mentales, su cuerpo y el entorno muy distinta de la que ejemplifica una persona que lo hace con la autoridad que en condiciones normales reconocemos y, en este sentido, las posibilidades de acción resultantes de dicho individuo, dados sus estados mentales y el entorno que lo rodea, se encuentran drásticamente alteradas y limitadas.¹²⁸

¹²⁸ Finkelstein apunta a la clave de la autoridad respecto a los propios mentales reside en la capacidad de expresarlos autoadscribiéndoselos uno mismo, algo que el paciente del terapeuta no puede hacer. Por una parte, consideramos que esta tesis todavía requeriría una explicación adicional de por qué este hecho explica lo que es importante y valioso en relación a la conciencia. Por otra parte, en nuestra reflexión, nos interesamos más por el marco general que permite entender estas cuestiones, y en este sentido, consideramos que el valor de nuestra experiencia consciente no se puede determinar de un modo general, al margen de unidades relacionales concretas en las que el sujeto se encuentra incardinado.

Si el dinamismo entre el individuo y el entorno se mantiene estable, según lo dicho, no es inteligible que la llamada fenomenología varíe, ya que el conjunto de relaciones que el individuo mantiene con su entorno es constitutivo de lo que llamamos fenomenología.¹²⁹ Lo que esto sugiere es que lo que hace diferentes a los distintos sentidos o formas de percepción sensorial no es que generen, como un elemento adicional, una fenomenología especial en cada caso, sino que sus características y limitaciones nos permiten relacionarnos con nuestro entorno de un modo determinado. Igualmente, la idea del valor, e incluso del placer, es difícilmente comprensible como una especie de elemento o ítem añadido extrínsecamente y, en este sentido, si la experiencia consciente consiste en una serie de hechos adicionales e intrínsecos separados de todo lo demás, como señala Chalmers, esta concepción no puede dar cuenta de lo que la hace tan importante. La idea de una fuente intrínseca de valor, entendida en un sentido metafísico y último, resulta ininteligible, en la medida en que la idea de valor se encuentra necesariamente ligada a la noción de un ser anclado al mundo de una manera particular.¹³⁰¹³¹

El tipo de análisis que aquí presento sigue la línea de las teorías enactivas de la percepción (véase Nöe 2004, 2009 y O'Regan 2011). El enfoque enactivo es

¹²⁹ A modo de ilustración y, en cierto modo, corroboración de esta idea, es interesante tener en cuenta el experimento perceptivo llevado a cabo por Paul Bach-y-Rita (1972) en el que se diseñó un dispositivo que llegó a constituir una visión protésica. Dicho dispositivo se componía de una cámara de vídeo sujeta a unas gafas que se conectaba mediante a un cable a una especie de tablón dotado de vibradores que se ponía en contacto con la espalda o el abdomen del sujeto. Dichos vibradores generaban ciertos patrones en función de la información visual transmitida desde la cámara, de manera que el sujeto podía distinguir varios objetos, agarrarlos e incluso llegaban a distinguir rostros de personas; si accidentalmente presionaban el botón del *zoom* de la cámara, expandiendo la imagen bruscamente, los participantes retrocedían y se cubrían la cara protegiéndose. La pregunta que genera este experimento es, ¿la visión protésica debe considerarse como percepción táctil o cómo percepción visual? Hay que tener en cuenta que el artefacto táctil es muy limitado para poder trasladar la resolución de una cámara de vídeo. Sin embargo, si se pudiera trasladar fielmente la información visual a un sistema de patrones de estimulación táctil, permitiéndonos interactuar con nuestro entorno del mismo modo que nos lo permite nuestra percepción visual, ¿diríamos que hay una fenomenología intrínseca a la visión que no se encuentra presente en la estimulación táctil?

¹³⁰ Pensemos, por ejemplo, en la situación de transmitir en qué consiste el placer sexual a un ser no erótico.

¹³¹ Quizá la fijación de la sociedad contemporánea por entender lo valioso en términos de propiedad y posesión, como algo petrificado y estático, tenga su reflejo en una metafísica atomista que sólo sea capaz de pensar lo valioso y lo placentero como algo episódico y misteriosamente intrínseco; es por ello que a una mentalidad contemporánea le resulta tan ajena una noción de felicidad dinámica, como la *eudaimonia* aristotélica, o como la idea de que sólo se puede decir si una vida ha sido feliz al final de la misma. Por otra parte, la incapacidad filosófica de una imagen para poder dar cuenta de lo valioso puede tener un reflejo práctico en la incapacidad de poder llegar a tener una vida plena y satisfactoria.

problemático (véase Block 2005, 2012; Prinz 2006; Briscoe 2008)¹³² y una defensa detallada del mismo no solo excede el propósito de este trabajo, sino que es independiente de la argumentación presentada en este capítulo, basada en la insuficiencia del detectivismo para explicar el valor de la experiencia consciente. Esta argumentación, no obstante, apoya indirectamente el punto de partida del enfoque enactivo. Por otra parte, las imágenes de las que se sirven los filósofos enactivistas ayudan a pensar de un modo diferente sobre el problema de la conciencia y, especialmente, son una fuente de intuición que contrarresta las imágenes de los argumentos presentados en el Capítulo 3.

Finalmente, si el marco en el que se plantea el debate sobre la experiencia consciente, a la vez que plantea el problema mismo se encuentra ligado a una concepción de la experiencia consciente que no nos permite entender su valor, ¿qué deberíamos esperar de un enfoque científico de la experiencia consciente? ¿No es posible una ciencia de la experiencia consciente que explique sus aspectos más relevantes del mismo modo que el electro magnetismo pudo ser explicado? En la siguiente sección, analizaré el esbozo que Chalmers lleva a cabo de los principios generales con los que debería cumplir una ciencia de la conciencia.

6.5. La ciencia de la conciencia

En este capítulo he argumentado que la posición de Chalmers no dispone de recursos para poder dar cuenta del valor que atribuimos a la experiencia consciente, en la medida en que concibe la experiencia consciente, en su aspecto fenoménico, como una especie de ítem con un carácter intrínseco. El problema general de este tipo de respuesta no es tanto su falsedad, como el tipo de insatisfacción filosófica que genera. Si de lo que se trata es de comprender ciertos rasgos de la experiencia

¹³² Las teorías enactivas tienden a funcionar mejor con ciertos ejemplos, como las sensaciones táctiles, que con otros como las sensaciones de color y tienen múltiples problemas para dar cuenta de las experiencias visuales. En segundo lugar, las teorías enactivistas tienen problemas para poder dar cuenta de los sueños y las alucinaciones, en los que hay experiencias perceptivas sin que el cuerpo se encuentre interactuando con el mundo de la manera normal. Finalmente, es difícil desde una posición enactivista argumentar en contra de la intuición de que todas nuestras experiencias podrían ser reproducidas artificialmente por medio de la estimulación directa del cerebro.

fenoménica, remitir al carácter intrínseco de la fenomenología sin explicar por qué ésta es importante, o en qué consiste su carácter intrínseco, a lo sumo traslada de lugar la perplejidad básica. Por otra parte, Chalmers no se limita a defender una posición dualista respecto a la experiencia consciente, sino que trata de determinar unos principios generales que fundamenten una ciencia empírica de la conciencia. Aunque, de entrada, este proyecto puede resultar contradictorio, dado que uno de los rasgos que hacen poco atractivos el dualismo es su incompatibilidad con la ciencia, Chalmers propone introducir una revisión tanto respecto al tipo de teoría científica que se puede llegar a tener respecto a la experiencia consciente, como en cuanto a cuál es el lugar que ésta debe ocupar dado el marco fisicista. En este sentido, el diagnóstico de Chalmers es que el fisicismo es estrictamente falso a pesar de que, como veremos, la diferencia entre su posición y la de un materialista respecto a la experiencia consciente puede llegar a ser bastante sutil, de manera que la incorporación de la experiencia consciente al corpus de la ciencia no supondría una amenaza ni para el conocimiento científico, ni para la corrección, en términos generales, del materialismo como una concepción aproximadamente adecuada de la realidad. A continuación, cuestionaré la viabilidad de la ciencia de la conciencia propuesta por Chalmers, tanto en un sentido general como en relación a las pretensiones de su proyecto. A continuación, trataré de explicar de un modo tentativo en qué sentido es posible una ciencia de la conciencia y cuáles podrían ser sus límites.

6.5.1. La ciencia de la conciencia

Si bien Chalmers concluye que las propiedades fenoménicas no pueden quedar fijadas por los hechos físicos del mundo, esto no es incompatible con la posibilidad de desarrollar una ciencia de la experiencia consciente. Según Chalmers, en la historia de la ciencia ha habido momentos en los que ha sido necesario llevar a cabo una ampliación del catálogo ontológico de las distintas ciencias. En algunos casos, las nuevas entidades incorporadas no pueden reducirse a otros conceptos; en otros casos, los hechos que involucran dichas entidades no pueden fijarse a partir de

hechos que no las incluyan. Uno de los ejemplos que Chalmers utiliza es la ampliación que tuvo lugar en la física para poder incorporar las leyes del electromagnetismo. Del mismo modo, en el caso de las propiedades fenoménicas de la experiencia, el resultado de los argumentos antirreduccionistas no tiene por qué desembocar en el misterio, sino que el catálogo ontológico de la ciencia podría ampliarse una vez más, de manera que puedan determinarse leyes psicofísicas que sirvan para fijar las condiciones bajo las cuales hay conciencia y de qué manera determinados tipos de propiedades fenoménicas se correlacionan con estados funcionales, realizados por propiedades neurológicas que, en última instancia, son reducibles (al menos en principio), a propiedades físicas. De este modo, si bien Chalmers renuncia a que una ciencia de la conciencia pueda establecer algo más que una correlación estricta entre la fenomenología y la psicología, éste ya sería un resultado notable, no muy distinto de la incorporación del electromagnetismo a la física.

Siguiendo este razonamiento, el principio básico de la ciencia de la conciencia es que hay coherencia entre la fenomenología y la cognición. De entrada, si bien los datos básicos de dicha ciencia son problemáticos, Chalmers propone que se tomen en serio los juicios fenoménicos sobre nuestra propia experiencia. Aceptando este punto de partida, Chalmers propone el principio de fiabilidad: “Cuando juzgo que estoy teniendo una sensación auditiva, normalmente estoy teniendo una sensación auditiva. Cuando pienso que acabo de experimentar un dolor, normalmente he experimentado un dolor” (Chalmers 1996: 218-19). Un principio complementario a este es el principio de la detectabilidad, el cual apunta a la capacidad de formar juicios fenoménicos cuando se da la experiencia consciente. Si retomamos la clasificación propuesta por Chalmers respecto a los juicios fenoménicos, este principio incluiría los juicios fenoménicos de segundo orden, los cuales se refieren a experiencias y, especialmente, a sensaciones. Sin embargo, el principio más importante, según Chalmers, es el que afirma la coherencia entre la conciencia fenoménica y la conciencia psicológica (*awareness*). Para entender este principio, hay que tener en cuenta la diferencia entre los dos conceptos de lo mental. En su forma general, puede enunciarse afirmando que siempre que un sujeto tenga

una experiencia, éste será consciente (*aware*) de los contenidos de dicha experiencia. En este sentido, alguien que perciba un libro tendrá determinada experiencia o conjunto de experiencias (en el sentido fenoménico), de tal modo que la percepción del libro, según la ciencia de la conciencia, estará acompañada de dichas experiencias, y viceversa (el principio es válido en las dos direcciones)¹³³.

El último principio de la conciencia al que se refiere Chalmers hace más concreta la relación entre conciencia psicológica y fenomenología, de manera que “varios rasgos *estructurales* de la conciencia se corresponden directamente con rasgos estructurales que se representan en la conciencia psicológica (*awareness*)” (Chalmers 1996: 223). La idea que yace detrás de este principio es que la experiencia consciente no es un bloque homogéneo, sino que está dotada de cierta estructura. Por ejemplo, en el campo visual se pueden apreciar distintas formas geométricas, así como patrones, figuras o dimensiones. Según Chalmers, la experiencia, en su sentido psicológico, refleja la estructura de la fenomenología, y viceversa. En este sentido, siempre que en el campo visual de un sujeto pueda distinguirse una mancha amarilla redonda, situada a una determinada distancia de otra mancha, por ejemplo, verde y con una forma cuadrada, estas discriminaciones tendrán un reflejo funcional, de manera que se corresponderán con discriminaciones cognitivas. Sin embargo, a pesar de que este principio parece decir algo más detallado sobre la relación entre conciencia y cognición, no implica, por una parte, ningún cambio en la práctica de la ciencia cognitiva; por otra parte, tampoco puede explicar por qué la experiencia consciente es importante y por qué es un objeto digno de ser estudiado por la ciencia. El principio de coherencia estructural tampoco llevaría la ciencia de la conciencia más allá de una tabla de correspondencias en la que, en muchos casos, la parte de la fenomenología sólo

¹³³ Chalmers apunta a algunas dificultades que puede entrañar este principio. Por ejemplo, respecto a la memoria, uno puede ser consciente (*aware*) de que el Everest es la montaña más alta del mundo y, en principio, este es un estado mental que se encuentra presente continuamente y de un modo estable. Sin embargo, no parece que a este estado cognitivo le acompañe un estado fenoménico. En este sentido, Chalmers restringe la coherencia entre los dos conceptos de conciencia a pensamientos ocurrentes, que no incluyen aquellos que necesitan ser, de algún modo, recuperados, como en el caso de la memoria.

podría ser especificada por referencia, en algunos casos, a la percepción de un proceso cognitivo o a la percepción de un objeto físico¹³⁴.

Para que una teoría de este tipo sea posible, es necesario comprometerse con ciertos supuestos que permitan hablar de coherencia, correspondencia o de coincidencia estructural. En los capítulos 2 y 5, respectivamente, propuse dos argumentos que, de funcionar, pondrían en serias dificultades la viabilidad de este proyecto. El primero de ellos trataba de comprender cómo es posible establecer correlaciones entre estados funcionales e instancias de propiedades fenoménicas; para ello, era necesario poder individuar la identidad de un estado mental a través de cualidades fenoménicas, a las cuales sólo se puede acceder desde el punto de vista de la primera persona, puesto que cualquier vínculo con un criterio público (verbal, de conducta o físico) sería empírico, nunca constitutivo¹³⁵. Si este argumento funciona, el proyecto de la ciencia de Chalmers no es viable, ya que no se pueden individuar instancias de propiedades mentales exclusivamente desde el punto de vista de la primera persona y, por ende, no es posible establecer una correlación entre propiedades públicamente observables y propiedades fenoménicas. El problema que plantea el segundo argumento es que para poder establecer el tipo de conclusión a la que apunta la argumentación de Chalmers, es necesario comprometerse con una concepción de la experiencia consciente que resulta ser ininteligible. La concepción a la que me refiero parte del supuesto de que la experiencia consciente tiene un carácter discreto, de modo que es posible individuar fragmentos separados de experiencia consciente, y a su vez asume una posición fuertemente realista respecto a los estados mentales conscientes. Lo que este argumento trataba de mostrar es que la conjunción de la concepción discreta de la experiencia junto con una posición realista es una concepción ininteligible, de

¹³⁴ Por ejemplo, ¿cómo podría nombrarse la experiencia fenoménica del sabor del queso azul o cómo podría explicarse el dolor típico de una quemadura sin referencia al queso azul o a una quemadura? Ésta es otra forma de mirar al problema que se planteó en el Capítulo 2 en relación a la forma de individuar propiedades fenoménicas.

¹³⁵ Precisamente, la ciencia de Chalmers aspira a proporcionar un conocimiento ampliativo y, por lo tanto, la relación entre un criterio que permita atribuir experiencia consciente y la presencia de ésta tiene que ser contingente.

manera que, según este argumento, la ciencia de la conciencia de Chalmers tampoco sería posible¹³⁶.

Otra cuestión que no puede dejarse de lado es que, si recordamos la paradoja del juicio fenoménico, la ciencia de la conciencia de Chalmers sería *sui generis* en otro aspecto: el objeto principal de tal ciencia carecería completamente de eficacia causal. En este sentido, la analogía con el electromagnetismo no puede sostenerse, ya que la postulación de entidades que van más allá de la mecánica se hace necesaria, entre otras cosas, en la medida en que es importante comprender ciertos fenómenos que interfieren claramente en la causalidad del mundo. Por otra parte, es difícil tratar de encontrar una analogía científica alternativa que nos permita entender las particularidades de la ciencia de la conciencia. El planteamiento del problema encierra ya cierto elemento paradójico, ya que por una parte se trata de un fenómeno del que parece que dependa el valor que tiene la vida e incluso el estatus de los seres que gozan del mismo; por otra parte, parece que la experiencia consciente no añade nada relevante al mundo o, al menos, no está muy claro qué diferencia introduce la experiencia consciente en el mundo.

Si nos preguntamos cuál sería la diferencia práctica o tecnológica a la que la ciencia de Chalmers daría lugar, no podemos apuntar a algo concreto. Es evidente que la práctica y la investigación en ciencia cognitiva no cambiarían un ápice en su carácter, ya que el aspecto fenoménico de lo mental está fuera del alcance de cualquier progreso que se haga en ciencia cognitiva. Quizá podría establecerse una especie de catálogo o tabla que correlacione determinados estados cognitivos con sus correspondientes correlatos fenoménicos, teniendo en cuenta los testimonios de distintos sujetos. Sin embargo, tal y como se mostró en el Capítulo 2, el carácter de estas correlaciones no podría ser meramente empírico y no estaría claro el estatuto de una tal ciencia de la conciencia.

¹³⁶ A menos que el formato discreto de la experiencia se entienda en un sentido laxo y no metafísico, pero en este sentido no nos estaríamos refiriendo a un concepto demasiado definido de conciencia que esté a la altura de las pretensiones científicas a las que aspira la teoría de Chalmers.

Hasta ahora, he analizado la posibilidad de la ciencia de la conciencia de Chalmers en un sentido general. En la siguiente sección, dejaré de lado los argumentos precedentes y analizaré en qué medida la ciencia de la conciencia de Chalmers es una solución a los problemas que él mismo plantea.

6.5.2. Dualismo naturalista y fisicismo

Como ya anticipé, a pesar de que la posición de Chalmers es dualista, se trata de un tipo de dualismo *sui generis*, en la medida en que no sólo se diferencia sutilmente de una posición típicamente fisicista, sino que en cierto sentido trata de ser fiel al espíritu materialista, articulando un punto de vista que convierte la experiencia consciente en un elemento asimilable por el corpus científico. Se trata de lo que el mismo Chalmers denomina un dualismo naturalista. De entrada, el mismo nombre de la posición connota cierta tensión, en la medida en que el hecho de que un dualismo tenga que ser calificado con el adjetivo de naturalista apunta a una revisión notable, bien de lo que debemos entender por dualismo, o bien por naturalismo. Veamos cómo presenta Chalmers su posición:

Para atrapar el espíritu de la visión que defiendo, la llamo *dualismo naturalista*. Es naturalista porque postula que todo es una consecuencia de una red formada por propiedades básicas y leyes, y porque es compatible con todos los resultados de la ciencia contemporánea. Y como con las teorías naturalistas en otros campos, esta visión permite que podamos *explicar* la conciencia en términos de leyes naturales básicas. No tiene que haber nada especialmente trascendental respecto a la conciencia; se trata sólo de otro fenómeno natural. Lo único que ha ocurrido es que nuestra imagen de la naturaleza se ha expandido. A veces “naturalismo” se entiende como sinónimo de “materialismo”, pero me parece que un compromiso con una comprensión naturalista del mundo puede sobrevivir al fracaso del materialismo. (Si un materialista lo duda, señalo al resto de este trabajo como evidencia.) Alguien podría encontrar cierta ironía en el nombre de la visión, pero lo que es más importante es que transmite el mensaje central: aceptar el dualismo no es necesariamente aceptar el misterio. (Chalmers 1996: 128)

Una de las posiciones de las que se tiene que desmarcar este tipo de dualismo es el dualismo cartesiano, o al menos una de las consecuencias derivada del mismo que resultaría incompatible con cualquier posición que, en un sentido mínimo, pueda considerarse fisicista, a saber, la falsedad de la clausura causal del mundo físico. De acuerdo con este principio, todo suceso físico tiene una causa completa antecedente física o, en su versión explicativa, todo suceso físico puede ser explicado en términos exclusivamente físicos. Según el dualismo cartesiano, este principio no se cumpliría, ya que la *res cogitans* interfiere causalmente en el mundo material. Cuando Chalmers habla de misterio, probablemente se esté refiriendo a la violación de este principio. Si la ciencia puede aceptar entidades no materiales en su repertorio ontológico, el único contenido que le resta a la noción de misterio es que los distintos fenómenos a comprender escapen a la posibilidad de ser predecibles o subsumidos, en principio, bajo ciertas leyes. Como ya se señaló en el Capítulo 1, el objetivo de evitar este tipo de consecuencia no es un rasgo exclusivo de una posición materialista. Un mundo en el que, por ejemplo, los objetos físicos se esfumaran y aparecieran constantemente, sin ningún tipo de regularidad, no sólo no sería un mundo en el que no sería posible descubrir leyes naturales, sino que quizá ni siquiera sería posible tener la noción de objeto, de manera que una ontología que no excluyera la posibilidad de dicho mundo sería rechazable por razones independientes de los argumentos fisicistas.

Por otra parte, la preocupación teórica relevante para el fisicista no puede ser que el mundo sea un conjunto de hechos encadenados arbitrariamente, ya que, en última instancia, si la experiencia es una fuente de conocimiento, en algún momento se tendrá que llegar a un lecho de conexiones brutas e inexplicables, que no puedan predecirse por la mera comprensión conceptual.¹³⁷ La inquietud del fisicismo no es tanto la existencia de arbitrariedad en el mundo, como la existencia de conexiones brutas que no se reduzcan, en última instancia, a las conexiones brutas que

¹³⁷ Esta dialéctica tiene sentido dentro del marco fisicista. Como ya vimos en el Capítulo 4, la idea de alcanzar un nivel último, en este caso de conexiones brutas, tiene sentido si se presupone una concepción estratificada de la realidad, supuesto altamente cuestionable que parece estar detrás de las intuiciones fisicistas más poderosas.

sustentan las leyes de la física. El movimiento teórico de Chalmers consiste en ampliar el conjunto de propiedades que pueden figurar en las leyes científicas, con el fin de cubrir el ámbito de la experiencia fenoménica, de la misma manera que la inclusión del electromagnetismo en la física supuso la postulación de ciertas entidades básicas que no podían reducirse a propiedades mecánicas. Sin embargo, la analogía no se sostiene de un modo evidente. La ampliación de la física en el caso del electromagnetismo no supuso una ruptura que, actualmente, se considere como el fracaso del materialismo; sin embargo, según Chalmers, los hechos relativos a la experiencia fenoménica son incompatibles con la verdad del materialismo. Podría decirse que, dado que las propiedades fenoménicas no figuran en las leyes de la física, la ampliación propuesta por Chalmers es incompatible con el fisicismo. Pero, ¿qué razones hay para rechazar esta ampliación como perteneciente todavía al ámbito de la física? Al fin y al cabo, teniendo en cuenta la incompletud de la física actual, es probable que en el futuro se amplíe el catálogo de su ontología o que incluso se revise profundamente. En el Capítulo 1 traté con mayor detalle esta cuestión y una de las conclusiones a las que llegué es que el contenido del fisicismo respecto a lo mental se ha minimizado en el debate actual hasta el punto de que lo que está en juego es que lo mental sea reducible a lo no mental. Ahora bien, ¿qué razones hay para considerar que hay una línea que separa lo mental y lo no mental y que resulta problemática? Uno de los desafíos que tiene que afrontar el fisicismo es cómo dar cuenta de la existencia de las propiedades semánticas del mundo.¹³⁸ Éste no es un problema que se circunscriba al ámbito de la experiencia fenoménica; de hecho, el problema de la experiencia fenoménica parece un problema de un carácter distinto, de modo que, si hubiera una teoría reduccionista exitosa respecto a los estados intencionales, no se seguiría trivialmente una reducción de las propiedades fenoménicas. La pregunta respecto a cómo mis creencias pueden ser sobre el mundo, de manera que tengan un contenido que describe un estado de cosas, es una pregunta distinta de cómo o

¹³⁸ Según Daniel Stoljar (2010), la gran tarea de la filosofía dentro de un marco fisicista consiste en redefinir o reinterpretar distintas esferas de la realidad, de manera que puedan mostrarse compatibles con la tesis fisicista. El significado es una de estas esferas, pero también se refiere al ámbito de la moral y a la modalidad (lo que él mismo llama las tres Ms: *Meaning, Moral and Modality*).

bajo qué condiciones se da la experiencia consciente. Esta última cuestión, así formulada, se encuentra mucho más cerca de la cuestión de formular un antecedente causal que garantice, por ejemplo, la generación de un campo magnético. En este sentido, la ampliación de Chalmers del discurso científico con el fin de englobar las experiencias fenoménicas supondría una respuesta que encaja mejor con este último tipo de cuestión. Sin embargo, como ya se señaló varias veces, el catálogo de lo que Chalmers considera ejemplos de estados mentales con un carácter fenoménico es muy heterogéneo, y dependiendo de cuál sea el ejemplo que se tenga en cuestión, las intuiciones respecto a si dicho estado debería incluirse en el ámbito de lo psicológico o de lo fenoménico varían notablemente. Chalmers es consciente de esta ambigüedad y señala que la clasificación de los estados mentales tiene un carácter convencional, sin vincularse a una cuestión filosófica profunda, aunque dicha ambigüedad deba aclararse en una discusión teórica. En el caso de una creencia, parece que el aspecto psicológico predomina, mientras que en el caso del dolor o de una experiencia olfativa hay un vínculo más claro con el ámbito de lo fenoménico. Sin embargo, esta advertencia metodológica es insuficiente, como mostraré a continuación.

Volviendo a la concepción fotográfica, una de sus consecuencias es que se genera una duplicidad de propiedades, de manera que a algunas propiedades fenoménicas les corresponde una propiedad en el mundo. Cuando alguien contempla un tomate, por ejemplo, habría que distinguir la propiedad fenoménica de rojo, del rojo como una cualidad propia del tomate.¹³⁹ En el caso de un olor, parece que esta distinción no puede llevarse a cabo de un modo tan claro, en la medida en que el olor es separable del objeto de un modo en el que un color no lo es. Por ejemplo, podemos percibir el olor de una naranja sin percibir la naranja, pero no podemos percibir el color de una naranja sin percibir la naranja.¹⁴⁰ Y si se trata de un dolor, esta distinción no puede trazarse, ya que ¿en qué podría consistir que un objeto es objetivamente doloroso, a no ser en la medida en que nos causa

¹³⁹ Ésta distinción equivale a la distinción entre “ver predicativo” y “ver de sensación” llevada a cabo en Stroud (1999).

¹⁴⁰ En este caso, claro está, no me refiero a percibir el color de la naranja como tipo, sino a la ejemplificación particular de color de la naranja.

sensaciones dolorosas? Por otra parte, un proceso puede ser doloroso, pero es difícil pensar en situaciones en las que podría tener un color. En el caso del color hay ejemplos que se entienden mejor con el modelo de la sensación, como en el caso del Arco Iris, en el que no existe ningún objeto de color o de una luz coloreada. ¿Y qué ocurre respecto a las percepciones de formas, puede hablarse de la propiedad fenoménica asociada a la contemplación de algo cuadrado, o a una propiedad fenoménica ovoide? Todos estos ejemplos muestran que el catálogo de Chalmers incluye elementos entre los que se pueden trazar distinciones que, si bien no eliminan o invalidan la formulación del problema de la conciencia, lo matizan y quizá lo fragmentan en una serie de problemas diferentes y, en cierto grado, independientes entre sí. En cualquier caso, si aceptamos que la concepción de la experiencia consciente de Chalmers es fotográfica, según los rasgos descritos, entonces la teoría de Chalmers se tiene que enfrentar a dos problemas generales: por una parte, tiene que ser capaz de fijar el contenido de las propiedades fenoménicas de la experiencia con independencia de criterios observables públicamente; por otra parte, tiene que explicar cuál es la relación entre las propiedades fenoménicas de la experiencia y las propiedades correspondientes atribuidas a objetos del mundo circundante. La imposibilidad de dar una solución al primer problema se mostró en el Capítulo 2; respecto al segundo problema, Chalmers no podrá dar una solución que satisfaga las pretensiones de su proyecto, como se muestra a continuación.

6.5.3. Propiedades fenoménicas

El dualismo naturalista de Chalmers se resume en las dos siguientes tesis: en primer lugar, el fisicismo es falso, de manera que los hechos relativos a las propiedades fenoménicas son adicionales respecto a los hechos físicos y no tienen una conexión lógica con estos últimos; en segundo lugar, hay una conexión nomológica entre los hechos relativos a la fenomenología y los hechos físicos, de manera que en nuestro mundo existen leyes naturales que conectan los hechos físicos con la fenomenología. Ahondar en el carácter de esta conexión se encuentra fuera de

lugar, ya que como vimos en el Capítulo 1, dentro del marco fisicista de Chalmers se aspira a mostrar que todos los hechos del mundo se derivan lógicamente de los hechos físicos, si bien las leyes últimas de la física deben ser aceptadas como hechos brutos. El fisicismo no es una teoría que trate de hacer inteligibles, en algún sentido, los hechos físicos, sino que argumenta que éstos son los hechos básicos del mundo.¹⁴¹ El proyecto de Chalmers, si bien es el resultado de las aspiraciones vinculadas a la concepción absoluta de la realidad, es más exigente en la medida en que parte de la posibilidad inicial de considerar los hechos físicos como hechos de los que se siguen lógicamente el resto de hechos. Además, podría pensarse que la mera correlación nomológica entre tipos de hechos no es suficiente para satisfacer los requisitos de la concepción absoluta de la realidad, ya que ésta impone un requisito explicativo que la mera correlación no parece satisfacer. Sin embargo, Chalmers define su modelo de explicación reductiva en términos de la correlación expresada por la noción de supervinencia. Si bien en ocasiones Chalmers señala de manera provisional que ambas nociones no deberían confundirse, no está muy claro qué margen queda en su planteamiento para definir otra noción de explicación.

Chalmers reconoce el fracaso del proyecto fisicista, pero a la vez considera que mantiene su espíritu al llevar a cabo una ampliación del catálogo de elementos que debería incluir la ciencia. Aunque Chalmers interpreta este movimiento a la luz de la ampliación que sufrió la física para incluir el electromagnetismo, este símil resulta paradójico en la medida en que la ampliación conceptual a las propiedades fenoménicas supone el reconocimiento de la falsedad del fisicismo. Por otra parte, la indefinición del contenido concreto que se asigna a lo físico en la discusión resta importancia a la cuestión de si el fisicismo es verdadero o falso. Ahora bien, si la conexión entre los hechos físicos y la existencia de propiedades fenoménicas no se da en otros mundos posibles, de un modo abstracto y negativo se responde a la

¹⁴¹ En este sentido, cabría puntualizar que hay diferencias de matiz entre distintas posiciones fisicistas. Según una posición microfisicista, que es la que parece aceptar Chalmers, dentro de las leyes de la física podría aceptarse cierta jerarquía, de manera que los hechos físicos son reducibles a las leyes que rigen los hechos microfísicos. Otros autores, como Andrew Melnyk, no aspiran a articular esta jerarquía respecto a los hechos físicos.

primera de las preguntas planteadas al final de la sección anterior. El hecho de que percibamos un objeto rojo se vincula contingentemente y de un modo bruto con el hecho de que tengamos una experiencia con la propiedad fenoménica o el *quale* de rojo. No obstante, ¿puede aceptarse, de entrada, que el vínculo entre nuestros estados mentales conscientes y las propiedades objetivas del mundo sea una correlación bruta? La paradoja del juicio fenoménico ilustra de un modo intuitivo la implausibilidad, *prima facie*, de la idea de que los juicios sobre nuestra propia experiencia se vinculen con nuestras experiencias de un modo bruto; sin embargo, el problema tratado aquí es más específico. Los dos aspectos de lo mental que distingue Chalmers generan una duplicidad, de manera que por cada estado mental consciente se puede distinguir un aspecto psicológico y uno fenoménico. Un zombi sólo dispondría de estados mentales con un aspecto psicológico. Sin embargo, un zombi sería un detector perfecto de propiedades fenoménicas o, al menos, de las condiciones en las que un ser humano normal tendría determinado tipo de fenomenología, en la medida en que esta última se encuentra correlacionada con el aspecto funcional de la mente. Una consecuencia destacable de esta correlación es que un zombi, desconociendo, por definición, el aspecto fenoménico del color, sería capaz de llevar a cabo las mismas discriminaciones de color, y con los mismos detalles, que lleva a cabo una persona normal. Por lo tanto, si los zombis son lógicamente posibles, entonces puede haber conjuntos de propiedades físicas en los objetos físicos que se correspondan con nuestras sensaciones de color. Del mismo modo, habrá propiedades en los objetos físicos que se correspondan con las propiedades fenoménicas implicadas en la experiencia de formas, olores, tacto, etc. El problema es que al ser esta correspondencia una correlación bruta, parece que deja de lado un tipo de inteligibilidad propia de los estados mentales, a saber, la posibilidad de que nuestros estados mentales refieran o representen aspectos del mundo. En este sentido, Barry Stroud ha apuntado a una limitación básica de este modelo:

La correlación, de este modo, no da una explicación de por qué el tipo de percepción correlacionada con ciertos movimientos físicos es una percepción de aquello de lo

que es percepción. ¿Por qué lo que surge precisamente de estas condiciones físicas es una percepción de amarillo? Eso es lo que queremos saber. Incluso si damos por sentado que *algo* psicológico surge precisamente de estas condiciones, ¿en este sentido, por qué no es una percepción de rojo, o una percepción de algo cúbico, o un dolor—o un pensamiento sobre la raíz cuadrada de menos dos? ¿Por qué, en efecto, surge aquí, precisamente, ese fenómeno psicológico particular? Eso es lo que no podríamos comprender ya, según esta imagen. Tendría que aceptarse, simplemente, que ciertos tipos de ítems físicos y psicológicos se correlacionan en el mundo tal y como es. (Stroud 2000: 92)

Esta objeción, no obstante, no resulta igualmente convincente en todos los casos descritos. Parece muy claro que la percepción de un cubo tiene el carácter cualitativo que tiene en virtud de la forma del cubo, y una explicación de la misma que se reduzca a un nexo bruto entre ciertas condiciones físicas y cierta experiencia parece incapaz de dar cuenta de esta intuición. Sin embargo, suele asumirse que no ocurre lo mismo con las experiencias cromáticas, los olores o los dolores. Como ya hemos visto a lo largo de los otros capítulos, las intuiciones utilizadas en los argumentos de Chalmers se apoyan con mayor fuerza en determinados ejemplos, aunque trate de extrapolar sus conclusiones a todo tipo de estado mental consciente. Esta asimetría dejaría a un lado, por ejemplo, cualidades de forma o de movimiento y, al otro, cualidades como los colores, los sabores o los olores, considerando que las primeras producen en nosotros sensaciones que representan adecuadamente rasgos del mundo, mientras que las segundas son puramente subjetivas, de manera que en un mundo sin seres vivos habría formas cúbicas, pero no colores. Me estoy refiriendo, claro está, a la clásica distinción filosófica entre cualidades primarias y cualidades secundarias. En el Capítulo 1 se habló de la concepción absoluta de la realidad, como un proyecto teórico que trata de alcanzar una descripción privilegiada de la realidad, en la medida en que representa el mundo tal y como es, y se estableció un vínculo entre dicho proyecto y el planteamiento del problema de la conciencia. En estos momentos, nos encontramos en condiciones de describir con mayor precisión este vínculo.

No trataré de una manera directa la cuestión de si es posible una concepción absoluta de la realidad, ni de si la distinción lockeana entre cualidades primarias y secundarias, vinculada a dicha concepción, es sostenible, ya que ambas cuestiones exceden el ámbito de esta discusión¹⁴². Sin embargo, podría decirse que, si el problema de la conciencia es inteligible y no puede solucionarse dentro de un marco fisicista, esto supone el fracaso de la concepción absoluta de la realidad. Aun si se reconoce que la experiencia consciente es una realidad meramente subjetiva, que no responde a cómo son las cosas con independencia de seres que tienen estados subjetivos, dichos estados igualmente existen, aunque su existencia sea *sui generis*. Y si la concepción absoluta de la realidad, en la medida en que pueda desarrollarse, no logra integrar la experiencia consciente, entonces hay que reconocer su fracaso, con independencia de que lo que reste sean propiedades subjetivas. Por otra parte, el planteamiento del problema de la conciencia sólo es inteligible en el marco de un intento de alcanzar una concepción absoluta de la realidad; el único lugar que puede ocupar un punto de vista en dicha concepción es, en cierto sentido, provisional, en la medida en que la descripción privilegiada que se llegaría a alcanzar explicaría los distintos puntos de vista de manera que éstos podrían comprenderse a partir de una concepción que describiría las cosas tal y como son; la experiencia fenoménica, dado este planteamiento, constituiría el residuo que no permite ser incluido en la descripción de las cosas tal y como son.

El problema de la correlación bruta entre estados físicos y estados mentales es más intuitiva aplicada a cualidades primarias, ya que la conexión entre nuestra experiencia de algo cuadrado y que de hecho percibamos un objeto cuadrado apunta claramente a algo más que una correlación bruta, en la medida en que parece haber una correspondencia no contingente entre nuestra percepción de cuadrado y la existencia de algo cuadrado que produce dicha percepción (incluso mantenemos el mismo predicado para el objeto y la percepción). Sin embargo, tampoco es éste un problema que deba afectar demasiado a Chalmers, ya que los dos aspectos de lo mental pueden explicar la diferencia entre los dos casos. En el

¹⁴² Barry Stroud (2000) argumenta en contra de la posibilidad de mantener, coherentemente, la distinción entre cualidades primarias y secundarias.

caso de la percepción de algo cuadrado, hay un sentido funcional según el cual el hecho de que haya delante de alguien un objeto cuadrado no se vincula arbitrariamente con una experiencia de algo cuadrado, o al menos no en el sentido de arbitrariedad que hay entre ciertas condiciones físicas y una experiencia de amarillo; la concepción absoluta de la realidad no puede renunciar, en última instancia, a la existencia de hechos brutos, aunque una vez determinadas las leyes naturales y los hechos físicos del mundo, ya no puede haber sorpresas. No obstante, todavía podría quedar un residuo de experiencia fenoménica vinculada de un modo bruto, en un sentido no permitido por la concepción absoluta de la realidad, a la percepción de algo cuadrado que hiciera naufragar el proyecto general.

En última instancia, la distinción entre cualidades primarias y secundarias, la distinción entre los dos conceptos de lo mental y la distinción entre intencionalidad y fenomenología mantienen en común la idea de que hay dos tipos de estados mentales cuya identidad viene dada por su forma especial de relacionarse con el mundo. Por una parte, hay estados que mantienen una relación representacional con el mundo, mientras que hay estados que mantienen una relación bruta con éste; el tipo de ciencia que propone Chalmers supone que estas conexiones brutas son el cemento del universo, y a lo sumo podemos registrar dichas conexiones y determinar las regularidades que siguen. Además, las relaciones representacionales pueden reducirse, en principio, a dichas conexiones brutas. Sin embargo, dejando de lado la viabilidad de este proyecto, parece que en el caso de los estados mentales se genera una tensión profunda cuando el vínculo entre estos y el resto del mundo se convierte en una conexión bruta. Y es importante insistir en que este es un problema independiente del proyecto fisicista ya que, aun si la intencionalidad se pudiera reducir a propiedades no intencionales, todavía habría que entender, al menos, por qué una percepción de un cubo en presencia de un cubo constituye un tipo de relación no bruta. De otro modo, siguiendo la argumentación de Barry Stroud, el *explanandum* desaparece, y no puede entenderse qué significa que parece que determinados estados mentales, como las sensaciones, se corresponden ilusoriamente con algo que está genuinamente en el mundo, si bien no hay unas propiedades tales en el mundo tal y como es en sí.

Precisamente, el problema de la conciencia se genera al entender que, si bien los estados fenoménicos parecen relacionarse de un modo bruto con el mundo, esto los convierte en un tipo de propiedad extraña, de un modo distinto a como lo serían las propiedades electromagnéticas respecto a las mecánicas, ya que la relación entre dichas propiedades no nos produce el mismo tipo de perplejidad que en el caso de la fenomenología y el resto del mundo. Esta tensión se percibe como una insatisfacción en el caso de Chalmers, ya que, si bien presenta el problema de la conciencia como el problema más difícil al que se enfrenta el fisicismo, su solución pasa por tratarlo como un problema más de física, donde claramente no se genera ese tipo de perplejidad.

En el modelo esbozado en las secciones anteriores no está claro que pueda entenderse la experiencia consciente de un modo que precipite en unas propiedades que, de algún modo, se vinculen de un modo bruto con el resto de hechos del mundo. Una alternativa a la concepción fotográfica de la experiencia que se tuvo en cuenta fue que las llamadas propiedades fenoménicas no son sino propiedades del mundo. En este sentido, según este modelo, ni siquiera sería posible trazar la distinción entre ambos tipos de propiedades. Ahora bien, esta no puede ser toda la historia, ya que los problemas relativos a la experiencia consciente parece que no son comparables a otros problemas: ¿hay algo que pueda explicar por qué en el caso de la experiencia consciente se genera esa perplejidad? ¿Y por qué esa perplejidad puede generarse planteando experimentos mentales? ¿No supone ésta una gran diferencia respecto a casos como el del electromagnetismo?

El juego entre estas dos maneras de aproximarse al problema es lo que le permite a Chalmers hablar de plausibilidad respecto a los principios de su ciencia:

Todo tipo de consideraciones de plausibilidad juegan un rol en la formación de nuestras teorías, más allá del rol desempeñado por la evidencia, en todo tipo de dominios. [...] La evidencia empírica no es todo lo que tenemos que seguir en la formación de teorías; también hay principios de plausibilidad, simplicidad y estética, entre otras consideraciones. (Chalmers 1996: 216)

En el caso de los principios de la ciencia de la conciencia que propone Chalmers, recurrir a la noción de plausibilidad es un movimiento que quizá no puede llevar a cabo de un modo coherente. Para poder hablar de explicación, es necesario poder hablar de casos en los que no hay explicación; puede haber una ausencia de explicación de un hecho concreto, pero también puede tratarse de la ausencia de explicación de un tipo de hechos. Respecto a los hechos que tienen que ver con la experiencia fenoménica, no sólo parece que no hay una explicación de un fenómeno, sino que no se puede establecer un modelo de explicación que sea aplicable a otro ámbito de la naturaleza. El problema difícil (*hard problema*) de la conciencia, en este sentido, no puede entenderse como una laguna más en el corpus del conocimiento científico, tal y como se compara frecuentemente con la explicación de la vida biológica, sino que ni siquiera existe una inteligibilidad previa de qué contaría como una explicación. Ahora bien, al plantear una explicación como la más plausible, se está presuponiendo ya un modelo de explicación a partir del cual pueden incluirse distintas teorías candidatas a ofrecer la mejor explicación. Pero si el problema difícil es tal, este es un movimiento que está vedado, ya que de entrada no existe un modelo que permita dicha explicación, de ahí el cariz conceptual con el que parece presentarse el problema de la conciencia.

En este sentido, tampoco resultan iluminadores los experimentos mentales que Chalmers diseña para tratar de contrarrestar los argumentos de los *qualia* ausentes y del espectro invertido, respectivamente. En dichos experimentos, Chalmers plantea ciertas posibilidades respecto a un par de individuos funcionalmente isomórficos. Chalmers propone el experimento mental de los *qualia* desvanecidos como un rechazo de la posibilidad empírica de que una organización funcional que consideramos que está siempre acompañada de conciencia pueda carecer completamente de ésta; esta es la posibilidad de los “*qualia* ausentes”. El argumento de Chalmers es una reducción al absurdo que asume la existencia de *qualia* ausentes. Partiendo de esta posibilidad, imaginemos que nuestras neuronas pudieran reemplazarse gradualmente por chips de silicio, manteniendo constante nuestra organización funcional; asumamos que si dicha sustitución fuese completa nos convertiríamos en un robot sin *qualia*; esta sería la posibilidad de los *qualia*

ausentes. Chalmers llama la atención sobre cómo interpretar los estadios intermedios de la sustitución y plantea el siguiente silogismo disyuntivo: o bien la conciencia se va desvaneciendo gradualmente, o bien hay un momento de la sustitución en el que los *qualia* desaparecen abruptamente. Chalmers rechaza rápidamente la segunda posibilidad, ya que esto implicaría que una sola neurona podría hacer que se desvaneciera la conciencia (o partes significativas de la misma), pero esto es algo completamente arbitrario y, por ende, implausible. Respecto a la segunda posibilidad, Chalmers imagina un estadio intermedio de la sustitución en la que un duplicado funcional nuestro se encuentra profiriendo las mismas oraciones y realizando las mismas discriminaciones perceptivas, a pesar de que experimenta cualidades que han perdido, gradualmente, parte de sus características (por ejemplo, el paso de un rojo brillante a un rosa pálido). No obstante, independientemente de que algunas de las propiedades fenoménicas de la experiencia de nuestro duplicado funcional desaparezcan gradualmente, llegará un momento en que la detección funcional de propiedades psicológicas no se corresponderá, en absoluto, con las propiedades fenoménicas que en ese momento se estén instanciando, lo cual, según Chalmers, es altamente implausible, como nos muestran los casos con los que estamos familiarizados. Sin embargo, ¿el problema de la conciencia no cuestiona, de entrada, el conocimiento que podamos tener de los casos familiares de conciencia? ¿No es éste un movimiento circular?

Chalmers recurre a otro experimento mental que trata de responder a la posibilidad del espectro invertido y que llama los “*qualia* danzantes”. En este caso, se juega con la posibilidad, de nuevo, de dos duplicados funcionales que, *ex hypothesi*, tienen un color del espectro cromático invertido. Una de las novedades del experimento es que uno de los duplicados tiene un pequeño circuito formado por chips de silicio, donde debería haber un conjunto de neuronas. El otro duplicado, que podría ser el lector de este trabajo, se encuentra conectado a un circuito de silicio idéntico al del duplicado, si bien este circuito sólo comenzará a funcionar si se activa un interruptor. Tal y como está construido el ejemplo, la conexión con el circuito de silicio permite cambiar los *qualia* de color del individuo conectado con el circuito de un modo brusco, sin que esto influya en sus propiedades funcionales, de

manera que las experiencias del individuo podrían cambiar en un instante de rojo a azul sin que éste se percate. Ahora bien, una vez más, esta disociación entre las propiedades psicológicas de la experiencia y la fenomenología es muy implausible.

Los experimentos mentales de Chalmers son, más bien, ilustraciones que tratan de mostrar la plausibilidad del principio de la no variación organizacional, pero no constituyen un argumento que preste un apoyo independiente a dicho principio, sino que más bien lo presupone. En este sentido, no son iluminadores respecto a las intuiciones que ya estaban en juego al inicio de la discusión, sino que más bien se sirven de éstas y apelan a una noción de plausibilidad que no es clara, ya que reposa en los casos familiares de atribución conciencia, los cuales deberían poder comprenderse con una teoría de la conciencia.

Chalmers, en otros fragmentos, es consciente de que lo que propone no puede ser llamado una explicación como tal:

Sigue siendo plausible que siempre que los sujetos tengan un tipo de proceso cerebral apropiado, tendrán un tipo de experiencia subjetiva asociado. Simplemente, necesitamos distinguir correlación de explicación. Incluso si los datos de la primera persona no se pueden explicar completamente en términos de datos de la tercera persona, los dos tipos de datos están fuertemente correlacionados. (Chalmers 2010: 40)

Como ya vimos en la sección anterior, los argumentos presentados en los Capítulos 2 y 5 plantean grandes dificultades a la posibilidad de que puedan establecerse correlaciones estrictas entre propiedades fenoménicas y estados funcionales. Ahora bien, si el proyecto de Chalmers no es viable, ¿cómo puede estudiarse entonces la conciencia? ¿Los enfoques de la neurología y la ciencia cognitiva dicen todo lo que puede decirse sobre la conciencia? Para poder responder esta cuestión, hay que llevar a cabo algunas consideraciones respecto al rol de la ciencia, en su conjunto, y cómo puede encajar el estudio de la conciencia en éste.

6.5.4. Ciencia y exhaustividad

Un supuesto que a menudo se toma como suelo firme en los debates de filosofía de la mente (si bien en filosofía de la ciencia es objeto de discusión) es el ideal de la unidad de la ciencia. La tesis de la unidad de la ciencia no es una tesis unívoca, pero según cierta concepción de la misma, la ciencia es una especie de actividad que genera principios y enunciados que se acumulan y que además constituyen un depósito de conocimiento tal, que en una situación ideal sigue produciendo y acumulando más enunciados verdaderos. Si bien es difícil defender que es así como funciona de hecho la ciencia, al menos podría ser éste un ideal regulativo.

Este ideal de unidad de la ciencia se encuentra conceptualmente e históricamente ligado a la concepción absoluta de la realidad; recordemos que la concepción absoluta de la realidad se concibe como el resultado al que llegaría, en general, el desarrollo del conocimiento. Si respecto a una determinada cuestión se encuentran dos explicaciones diferentes, ¿no debería poder darse cuenta de dicha diferencia? ¿No debería haber un metadiscurso que pudiera explicar por qué hay distintas explicaciones de un mismo fenómeno (o bien un metadiscurso que lleve a rechazar una de las explicaciones como inválida)? De otro modo, ¿sería inteligible la empresa del conocimiento?

En el Capítulo 2, indiqué que el problema de la conciencia se genera en el marco de una concepción absoluta de la realidad, y el tipo de ciencia de la conciencia que propone Chalmers es, precisamente, un intento de encontrar una perspectiva que abarque todos los hechos, de manera que la fijación de unos hechos básicos del mundo implica la fijación de los hechos del mundo *tout court*. Un criterio fundamental para medir el éxito del proyecto de Chalmers, desde su propio punto de vista, es la exhaustividad. El fisicismo no es una posición satisfactoria en la medida en que no es exhaustivo, de ahí que Chalmers perciba su dualismo naturalista como una posición superior, ya que, ampliando la ontología del fisicismo, logra fijar la totalidad de hechos del mundo.

Por otra parte, la noción de plausibilidad a la que recurre Chalmers también se nutre del supuesto de que es posible la unidad de la ciencia; en la medida en que es posible subsumir un conjunto de hechos bajo otros principios. Sin embargo, si bien el problema de la conciencia se genera en el seno de una concepción absoluta

de la realidad, la solución que propone Chalmers entra en conflicto con ésta y genera una tensión entre la concepción absoluta de la realidad y la unidad de la ciencia. La tesis de la unidad de la ciencia puede interpretarse de distintas maneras ¹⁴³, pero, en términos generales, podemos distinguir dos formas de concebir esta unidad: por una parte, la ciencia podría constituir una unidad internamente, de manera que las distintas ciencias se conectan entre sí mediante un determinado tipo de relación conceptual, que no necesariamente es la reducción. Por otra parte, las distintas ciencias podrían constituir una unidad en la medida en que hay un criterio externo que demarca la ciencia de lo que no lo es; este criterio podría ser un criterio metodológico.

La estrategia de Chalmers oscila entre las distintas problemáticas que plantean estos dos sentidos de unidad. Los argumentos antirreduccionistas de Chalmers se sitúan en el plano de las relaciones entre las distintas ciencias y el diagnóstico de Chalmers es que una ciencia de la conciencia no podría integrarse en la unidad que constituyen las distintas ciencias. Sin embargo, esto no impide que, metodológicamente, el estudio de la experiencia consciente no se diferencie del estudio de cualquier otro aspecto de la realidad, pero esta respuesta no es, claro está, una tesis reduccionista. Sin embargo, aun si dejamos de lado los argumentos de las secciones anteriores, cabría preguntar si la ciencia de la conciencia de Chalmers podría decirnos todo lo que es interesante respecto a la conciencia.

La ciencia de la conciencia de Chalmers parte del supuesto de que la conciencia es un fenómeno que puede estudiarse del mismo modo que se estudia el electromagnetismo. Sin embargo, tal y como hemos visto en los argumentos de este capítulo y del anterior, es difícil sostener que la experiencia consciente sea una “cosa” que pueda ser estudiada de un modo separado y aislado. Si esto es así, la ciencia que propone Chalmers, sin ser necesariamente falsa en algún sentido, no puede gozar del carácter explicativo de la física y la analogía con la ampliación del electromagnetismo no se sostendría.

¹⁴³ Véase Cat (2007).

En este capítulo, he tratado de mostrar la insuficiencia de la concepción de la mente de Chalmers para dar cuenta del valor de la experiencia consciente. Inicialmente, presenté una línea argumentativa continuadora de la crítica del Teatro Cartesiano que apuntaba a la ininteligibilidad de que los *qualia* constituyeran una fuente de valor, y señalé que desde una posición como la de Chalmers siempre se podía realizar un movimiento que reformulara el problema de la conciencia, ya que siempre es posible preguntar por qué determinado proceso o conjunto de propiedades está acompañado de experiencia consciente. A continuación, centré la reflexión en algunos casos de mentalidad inconsciente, con el fin de determinar cuál es el rasgo distintivo de los estados mentales conscientes. El sujeto tiende a tener un acceso privilegiado a sus propios estados mentales, de modo que es la mejor persona a la que preguntar para conocer sus estados mentales. Sin embargo, hay casos en los que el sujeto mantiene una relación anómala respecto a sus propios estados mentales que se mantienen en un nivel inconsciente; en estos casos, el sujeto puede tener una relación con su propia mente semejante a la que terceras personas tendrían con ésta. Este análisis llevó a la concepción de la conciencia que llamé detectivismo. Según el detectivismo, ser consciente de un estado mental se reduce a llegar a saber de él a través de un mecanismo de detección. Seguidamente, argumenté que, en todas sus formas, el detectivismo no puede dar cuenta de la autoridad de la primera persona ni del valor que atribuimos a la experiencia consciente, ya que la mera detección es la relación que una tercera persona podría tener con nuestros propios estados mentales y por lo tanto no puede explicar el valor que tiene la propia experiencia mental. Este diagnóstico era aplicable a la concepción de la mente de Chalmers, en la medida en que compartía los rasgos relevantes del detectivismo; para Chalmers, hay experiencia fenoménica cuando se ejemplifican ciertas propiedades, de manera que la experiencia consciente es la mera presencia de ciertos ítems cuyo carácter, *per se*, explicaría el valor que le atribuimos a estar conscientes. Seguidamente, presenté un marco alternativo para la comprensión de la experiencia consciente y el valor que le atribuimos que considera la fenomenología como estando constituida por el complejo haz de relaciones que el individuo mantiene con su entorno. No he

argumentado detalladamente a favor de esta concepción, sin embargo, la posibilidad de mirar al problema de la conciencia de una forma alternativa muestra que el dualismo naturalista no es la única concepción posible de la experiencia consciente y la dialéctica de Chalmers no es una senda forzada por una necesidad conceptual. Por otra parte, si bien las teorías enactivistas de la percepción son bastante controvertidas, gozan de un mayor potencial explicativo respecto al valor de la experiencia consciente. Finalmente, sostuve que el carácter explicativo de la ciencia de la conciencia que propone Chalmers no es comparable al de la física, de manera que no puede considerarse una continuación del naturalismo. Por otra parte, dicha ciencia se basa en el principio de que la conciencia es una “cosa” a ser estudiada separadamente, del mismo modo que se estudia cualquier otro fenómeno; sin embargo, esta es una posibilidad en contra de la que se ha argumentado en los capítulos precedentes.

Conclusions (Conclusiones)

This dissertation achieved three general results. Firstly, the assumptions that enable the formulation of the hard problem of consciousness are made explicit. Secondly, it is shown that Chalmers' account on consciousness cannot conclude intelligibly that consciousness is conceptually independent of public criteria of observation. Conscious mental states cannot be individuated without reference to objective processes and the assumption that conscious experience has a discrete metaphysical character, along with the realistic position Chalmers holds about mind, is unintelligible. Thirdly, a criticism is made on the kind of theory Chalmers develops, to the extent that it is unable to explain what is important regarding consciousness.

Physicalism, the philosophical approach that holds that physical facts fix all the facts of the world, is the framework within which the hard problem of consciousness is posed. Besides, the hard problem of consciousness seems to be an unavoidable consequence if physicalism is to be accepted. Is one of the conclusions of this dissertation the falsity of physicalism? To answer this question, it is important to take into account that no definite content of "physical" is essential for a physicalist position to be stated. The intuition that lies behind Chalmers' idea of physicalism is the metaphysical layered conception of reality. The importance attached to physics in the description of reality has to do with the location of physical properties in this hierarchical conception: physical properties seem to occupy an allegedly bottom-level and to the extent that physical properties do not occupy this bottom-level, physicalism would be false. In Chapter 4, I tried to show that intrinsic properties are such in relation to a context of normality within which some explanations make sense, while other ones are not explanations at all. If this argument is right, the further assumption behind this physicalist approach, namely, that a privileged description of reality can be given, to which every other perspective could be, at least in principle, reduced, is unintelligible. The intelligibility of explanations depends on what we take as a context of normality.

On the other hand, there are some uncontroversial facts that feed physicalist intuitions. These facts include the success and explanatory power of natural science,

the corroboration of natural laws in many practical contexts or the falsity of miracles. Regarding the mind, it is undeniable that there is no mind without brain and even a small change in the brain would have drastic and perhaps predictable consequences for the mind. However, the truth of these facts is logically independent of the truth of physicalism. None of the arguments presented goes against these well-established facts and I don't think any of the arguments of this dissertation has shown the falsity of physicalism either. The conclusion is rather that the theoretical project of physicalism cannot be carried out in an intelligible way. Additionally, the explanatory expectations of physicalism cannot be met because of the metaphysical constraints this very project puts on the way we think about the mind. The mind is thought of as a set of discrete states that exist independently of the possibility of knowing them. From a common-sense point of view there's nothing wrong with this way of describing the mind. Pains start at some moment and have an end at a concrete time. Moreover, we could be systematically wrong about somebody else's pain. Notwithstanding, problems pop up when this way of thinking about mental states is endowed with a metaphysical status, as I've tried to show in Chapters 5 and 6. In the same way as the Augustinian conception of language lacks explanatory power but cannot be said to be false, physicalism is limited by its metaphysical commitments to give a satisfactory account of reality.

On the other hand, if there is a continuity between the explanatory power of natural science and physicalism, a question arises regarding the possibility of dealing with consciousness as a further natural phenomenon. It is almost a commonplace in the discussion about consciousness that certain concepts can receive a scientific treatment and may be even as good objects for a physicalist reduction as far less controversial candidates for such a reduction. According to Chalmers, the psychological concept of the mental is perfectly amenable for this kind of treatment, in contrast with phenomenal consciousness, which should be recognised as an irreducible phenomenon.

Of course, saying that there cannot be a science of consciousness because there is no univocal concept of it is a superficial position. Such a conclusion could be countered with a list of the different types of concepts of consciousness and an

analysis of each of them that helps us determine which of those concepts are amenable to scientific treatment and which of them are out of our reach or even unintelligible. In the previous chapters, I have argued that the idea of phenomenal experience, as Chalmers understands it, fails to make sense. But then, what's the role of science in our understanding of consciousness? Does this mean that there is no genuine philosophical problem about consciousness? Perhaps, this latter response is also too hasty.

The problem of consciousness is at a crossroads where various philosophical problems converge, insofar as it takes into account, in its most general form, our subjectivity and the problems that our subjectivity and its relation to the world raise. The hard problem of consciousness is easy to formulate and the complex metaphysical tools that try to explain the intuitions at stake presuppose these intuitions instead of explaining them. What Chalmers dubs the hard problem of consciousness is already formulated in the writings of Locke and Leibniz. But does it make sense to talk about objectivity and subjectivity in such a general level of abstraction?

In last chapter, I suggested that the properties usually attributed to phenomenology are indeed properties of the world, perceived by an embodied and spatio-temporally located observer, endowed with specific psychomotor skills and interests. Then, the so-called phenomenology would be constituted by the set of interests and relations that an embodied individual maintains with her environment, so that understanding certain type of phenomenology would involve, in turn, understanding the network of relationships and interests that an individual endowed with certain attributes has with the world around him. But if this is so, the question of why consciousness arises from a series of neurophysiological processes cannot receive a general answer. Why a certain physiological process is accompanied by a red experience, rather than by a green experience? Can we answer this question without having some idea of how each colour relates to other colours, and how they vary depending on the positions an observer adopts in different environmental conditions? If the answer to this last question is negative, there is no substantive

problem of consciousness independently of a minimal knowledge regarding and embodied subject and the relations she holds to a determinate environment.

The content of Chalmers' anti-reductionist arguments cannot be determined independently of hypothetical scenarios that generate the intuition that a certain type of experience wouldn't exist, even if the conditions under which that experience usually obtains are met. According to Chalmers, those conditions are met by the functional organization of a given system. However, how is it that we can individuate a functional property as the correlate of a certain type of phenomenology is a question that Chalmers's account cannot properly address. Of course, there is a sense according to which talking about correlates of phenomenology is intelligible. It is clear that there are certain areas of the brain that are involved in vision, and even correlations can be established between small changes in the brain (as in the case of brain injury) and the experiences an individual has. The problem arises when this correlation is given some kind of absolute metaphysical character, as if by establishing a brute link between functional and phenomenal properties everything worth saying about consciousness would have been said. The kind of correlation Chalmers aims to establish is possible given a determinate whole that is taken for granted and, as in the case of the everyday concept of causation, the establishment of a correlation already implies an interest that defines a way of looking at things.

In this sense, whenever we try to explain why we experience things the way we do we will always find limitations that have to do with how our interests connect us to the world. For example, taking the example of Nagel, being a bat not only involves certain ways of perceiving the world, but also a complete lack of certain interests, including narrative interests. Even science is a form of storytelling that is motivated by certain interests. However, to know what it is like to be a bat without being so, it would be necessary to resort to some kind of narrative; but what could it be the narrative of a bat? The question is not that we cannot know what it is like to be a bat; the question is neither, following Wittgenstein's famous fragment, that we would not understand the speech of a bat about its experience. There is an even more fundamental problem, namely, that we cannot understand the possibility that a bat talks about its own experience, albeit in an unintelligible language. Given what

we know about bats there is no room in their lives for the slightest narrative interest, because in the life of a bat there is no way of looking at the world that can produce a narrative. While it is clear that a bat has some experience and, in some sense, knowing what it is like to be a bat goes beyond us, the experience of a bat cannot form a meaningful whole. If the view outlined regarding experience in the previous chapter is correct, the wealth of phenomenology is linked to the possibilities of action of a particular individual, and in this sense, the life of a bat is quite simple. Our experience is deeply modulated by the possibility of comparing different experiences, establishing relationships, remembering the past or anticipating the future, and how certain aspects of the world become visible depends on the ways in which our intelligence, our context and, of course, language, allow us to relate to the world.

Somehow, Thomas Nagel glimpses and comes to describe the character of this limitation, but fails to provide an explanation of it:

Our experience provides the basic material for our imagination, whose spectrum is thus limited. Not help trying to figure that has a membrane in the arms themselves, which allows flutter at dusk and dawn catching insects in one's mouth; who has poor vision and the surrounding world is perceived through a system of high-frequency sound signals reflected; and spent the day hanging upside down by his feet in an attic. To the extent we can imagine this (as you will not get far), it tells me what it would be for me to behave as a bat behaves. But that's not the point. I want to know how it is for a bat to be a bat. However, if I try to imagine this, I am restricted by the resources of my own mind, and those resources are inadequate to the task. I cannot carry it out or imagining additions to my present experience, or by imagining segments gradually subtracted from it, or by imagining some combination of additions, subtractions and modifications.

To the extent that I can look like and behave like a wasp or a bat without changing my fundamental structure, my experiences would not be at all like the experiences of those animals. (Nagel 1974: 439)

This fragment raises implicitly several questions. First, why the material of our imagination, as Nagel says, does not allow us to adequately represent what it is like to be a bat? Second, what is the difference between wondering what it is like, for me, being a bat and what it is like for a bat to be a bat? Thirdly, why fragments from my experience or additions to it cannot help me to know what it is like to be a bat? Finally, what is Nagel referring to when he speaks of fundamental structure?

If we consider what has been said in the last chapter, we can understand what kind of limitation Nagel pointed to. A priori, there is no reason why our imagination may be limited; however, by raising the second and third question, it seems it can be explained why it is not possible that we represent to us the inner life of a form of life relevantly different from ours. This is because an experience *for me*, to the extent that it is integrated in a complex framework that can become a narrative, only makes sense from the point of view of a physical being that is anchored to the world in a particular way. Given that we cannot relate to the world the way a bat does, because we do not possess the body of a bat, our psychomotor skills are very different and there are also many additional items that strongly affect the way we experience things (including those interests that characterize a human being), we cannot know what it is like to be a bat. On the other hand, when Nagel speaks of a "fundamental structure", he is not probably referring to something very definite or to some kind of essence. Rather, the chosen example is a sample of a different fundamental structure, insofar as psychomotor differences between a bat and a man are so evident that it is not even necessary to have too much knowledge about bats to generate a deep perplexity regarding how they experience the world.

The examples in the literature on conscious experience often appeal to bizarre possibilities, but the problem of consciousness could be raised with regard to many human experiences that are not available to everyone. In fact, it is not so difficult to find in any biography some experience that is alien to most people. Literature can, in this regard, help us to understand distant experiences insofar as it is able to create units of meaning that overlap with our own experiences and our own narratives. This is what we try to do by using pictures, metaphors or comparisons.

Of course, I am not denying that science can study consciousness, so that progress can be made in understanding how certain neurophysiological processes relate to certain experiences, and even this knowledge could be developed in a perhaps limitless degree of detail. The error is rather to assume that everything that can be said about consciousness is exhausted by the discourse of science and to consider conscious experience as an object of study that is accessible from a disembodied and totally detached eye. Consciousness is something too general to be exhausted by a single speech and it is, at the same time, an important and fundamental dimension of human life. If we consider that to understand the character of a particular experience it is necessary to understand how a particular being is related to her environment, the content of experience cannot be determined independently of the world. The philosophy of conscious experience has assumed experience has an intrinsic character and has ended up completely disassociating body and world. However, conscious experience cannot be separated from a determinate body and from concrete interests that constitute a particular way of being in the world.

Given what has been said, it is clear why the model of science of consciousness proposed by Chalmers does not allow to explain why there is consciousness at all. A further consequence of this proposal is that, even if Chalmers is right, no changes would be introduced in practices regarding the study of consciousness (and even if consciousness, in its phenomenological sense, were not to exist, these practices wouldn't change either). The type of brute link Chalmers establishes between phenomenology and the world cannot be understood by analogy with the ontological extension of other sciences, and Chalmers' science of consciousness does not meet the standard that very science should reach, according to him. Finally, I conclude that there cannot be a general science that covers all aspects of consciousness, to the extent that the so called problem of consciousness is not a general phenomenon that can be exhaustively dealt with by the current scientific research programs. The study of consciousness should rather yield an understanding of different forms of being conscious of different things, depending on how concrete embodied individuals are able to relate to their

environment and this understanding cannot be achieved without empirical knowledge. Hence, an appropriate approach to consciousness, rather than addressing a general question like the one expressed in the hard problem, should deal with many different questions in a multidisciplinary way. According to the proposal of this dissertation, the idea of a once-and-for-all established discourse (scientific or not) that says everything interesting about consciousness does not make sense.

Bibliografía

Armstrong, D., 1968. *A Materialist Theory of Mind*, London: Routledge and Kegan Paul.

Armstrong, D., 1980. "The Nature of Mind" en N. Block, ed., *Readings in the Philosophy of Psychology*, Vol I. Cambridge, Mass: Harvard.

Austin, John L., 1962. *Sense and Sensibilia*, London: Oxford University Press.

Baars, B., *A Cognitive Theory of Consciousness*, Cambridge: Cambridge University Press.

Bach-y-Rita, P., 1972. *Brain Mechanisms in Sensory Substitution*, New York and London: Academic Press.

Block, N., 1978. "Troubles with functionalism" en D. Chalmers, ed., *Philosophy of Mind: Classical and Contemporary Readings*. New York: Oxford University Press.

Block, N., 1990. "Inverted Earth." *Philosophical Perspectives*, 4, Tomberlin, ed., Atascadero, CA: Ridgeview.

Block, N., 2003. "Do causal powers drain away?", *Philosophy and Phenomenological Research*, 67, 133-50.

Block, N., 2005. "Review of *Action in Perception*", *Journal of Philosophy*, 102: 259–272.

Block, N., 2012. "Discussion of J. Kevin O'Regan's *Why Red Doesn't Sound Like a Bell: Understanding the Feel of Consciousness*", *Review of Philosophy and Psychology*, 3: 89–108.

Block, N. y Stalnaker, R., 1999. "Conceptual analysis, dualism, and the explanatory gap", *Philosophical Review*, 108/1: 1-46.

Briscoe, R., 2008. "Vision, Action, and Make-Perceive", *Mind and Language*, 23: 457–497.

Carnap, R., 1932. "Psychology in Physical Language", *Erkenntnis*, 3: 107-42.

Cat, J., 2013. "The Unity of Science", *The Stanford Encyclopedia of Philosophy* (Winter 2014 Edition), Edward N. Zalta (ed.), URL= <<http://plato.stanford.edu/archives/win2014/entries/scientific-unity/>>.

Chalmers, D. J., 1996. *The Conscious Mind*, Oxford: Oxford University Press.

Chalmers, D. J., 2000. "What is a Neural Correlate of Consciousness?" en T. Metzinger, ed., *Neural Correlates of Consciousness: Empirical and Conceptual Questions*. MIT Press.

Chalmers, D. J., 2002, "Consciousness and its Place in Nature" en D. Chalmers, ed., *Philosophy of Mind: Classical and Contemporary Readings*. New York: Oxford University Press.

Chalmers, D. J., 2002b. "Does Conceivability Entail Possibility?" en T. Gendler y J. Hawthorne, ed., *Conceivability and Possibility*. Oxford: Oxford University Press.

Chalmers, D. J., 2010. *The Character of Consciousness*. New York: Oxford University Press.

Churchland, P., 1981. "Eliminative materialism and the propositional attitudes", *The Journal of Philosophy*, 78/2: 67-90.

Corbí, J. E. y Prades, J. L., 2000. *Minds, Causes and Mechanisms: a Case Against Physicalism*, Oxford: Blackwell.

Crick, F., 1994. *The Astonishing Hypothesis: the Scientific Search for the Soul*. New York: Scribners.

Damasio, A., 1994: *Descartes' Error: Emotion, Reason and the Human Brain*. New York: G. P. Putnam's Sons.

Damasio, A., 2010. *Self Comes to Mind*, New York: Pantheon Books.

Davidson, D., 1970. "Mental Events" en D. Davidson, *Essays on Actions and Events*, Oxford: Oxford University Press, 207-223.

Dennett, D., 1991. *Consciousness Explained*, Boston: Little, Brown and Company.

Dennett, D., 2005. *Sweet Dreams: Philosophical Obstacles to a Science of Consciousness*, Cambridge, MA: The MIT Press.

Dupré, J., 1993. *The Disorder of Things: Metaphysical Foundations of the Disunity of Science*, Cambridge, MA: Harvard University Press.

Finkelstein, D. H., 2003. *Expression and the Inner*, Cambridge, MA: Harvard University Press.

Foster, J., 1991. *The Immaterial Self: A Defence of the Cartesian Dualist Conception of the Mind*, London: Routledge.

Goodman, N., 1978. *Ways of Worldmaking*, Hassocks, Sussex: Harvester.

Hardin, C. L., 1988. *Color for Philosophers: Unweaving the Rainbow*, Indianapolis: Hackett.

Harrison, B., 1973. *Form and Content*, Oxford: Blackwell.

Hill, C. S., 1997. "Imaginability, Conceivability, Possibility, and the Mind-body Problem", *Philosophical Studies*, 87: 61–85.

Hütteman, A., 2004. *What's Wrong with Microphysicalism?*, London: Routledge.

Huxley, T. H., 1874. "On the Hypothesis that Animals are Automata and its History" en D. Chalmers, ed., *Philosophy of Mind: Classical and Contemporary Readings*. New York: Oxford University Press.

Jackson, F., 1982. "Epiphenomenal qualia", *Philosophical Quarterly* 32, 127-36.

James, W., 1890. *The Principles of Psychology*, New York: Henry Holt and Company.

Kaplan, D., 1989. "Demonstratives" en *Themes from Kaplan*, Oxford: Oxford University Press, pp. 481-563.

Kim, J., 1998. *Mind in a Physical World*, Cambridge: Cambridge University Press.

Kim, J., 2005. *Physicalism. Or Something Near Enough*, Princeton: Princeton University Press.

Kolers P. A. y von Grünau, M. 1976. "Shape and color in apparent motion" en *Vision Research*, 16, pp. 329-335.

Korb. K. B., 1993. "Stage effects in the Cartesian Theater: a review of Daniel Dennett's *Consciousness Explained*", *Psyche*, 1(4).

Kripke, S., 1980. *Naming and Necessity*, Cambridge, MA: Harvard University Press.

Lewis, D., 1988. "What experience teaches" en D. Chalmers, ed., *Philosophy of Mind: Classical and Contemporary Readings*. New York: Oxford University Press.

Levine, J., 2001. *Purple Haze: The Puzzle of Conscious Experience*, Cambridge, Mass: the MIT Press.

Loar, B., 1990/1997. Phenomenal States. *Philosophical Perspectives* 4: 81–108. Revised en N. Block, O. Flanagan, y G. Güzeldere, ed., *The Nature of Consciousness* (Cambridge, Mass.: MIT Press, 1997).

McDowell, J., 1994. *Mind and World*, Cambridge, MA: Harvard University Press.

- McGinn, C., 1989. "Can we solve the mind-body problem?", *Mind* 98, 349-66.
- Mellor, D. H., 1974. 'In Defence of Dispositions', *The Philosophical Review*, 83: 157–181.
- Melnyk, A., 2003. *A Physicalist Manifesto: Thoroughly Modern Materialism*, Cambridge: Cambridge University.
- Moya, C., 2004. *Filosofía de la Mente*, Valencia: Universitat de València.
- Nagel, T., 1974. "What is it like to be a bat?", *Philosophical Review*, 4, 435-50.
- Nagel, T., 1983. *The View from Nowhere*, New York: Oxford.
- Nöe, A., 2004. *Action in Perception*, Cambridge, MA: Harvard University Press.
- Nöe, A., 2009. *Out of Our Heads*, New York: Hill and Wang.
- O'Regan, J.K., 2011. *Why Red Doesn't Sound Like a Bell: Explaining the Feel of Consciousness*, Oxford: Oxford University Press.
- Papineau, D., 1993. "Physicalism, Consciousness, and the Antipathetic Fallacy", *Australasian Journal of Philosophy*, 71: 169–83.
- Papineau, D., 2002. *Thinking about Consciousness*, Oxford: Oxford University Press.
- Poland, J., 1994. *Physicalism: The Philosophical Foundations*, Oxford: Clarendon.
- Popper, K. y Eccles, J., 1977. *The Self and Its Brain: An Argument for Interactionism*, New York: Springer.
- Prinz, J. J., 2006. "Putting the Brakes on Enactive Perception", *Psyche*, 12: 1–19.
- Putnam, H., 1960. "Minds and Machines" en *Dimensions of Mind*, S. Hook ed., New York: New York University Press.
- Putnam, H., 1975. "The Meaning of "Meaning"", en *Mind, Language, and Reality*. New York: Cambridge University Press.
- Rey, G., 1983. "A Reason for Doubting the Existence of Consciousness" en R. Davidson, G., Schwartz and D. Shapiro ed., *Consciousness and Self-Regulation* Vol 3. New York: Plenum.
- Rey, G., 1995. Toward a Projectivist Account of Conscious Experience, en T. Metzinger, ed., *Conscious Experience*. Lawrence, Kans.: Allen.

- Russell, B., 1912. *The Problems of Philosophy*, Oxford: Oxford University Press.
- Russell, B., 1927. *The Analysis of Matter*, London: Kegan Paul.
- Ryle, G., 1949. *The Concept of Mind*, London: Hutchinson.
- Schafer, J., 2003. "Is There a Fundamental Level?", *Noûs*, 37: 498-517.
- Schroeter, L., 2012. "Two-Dimensional Semantics", *The Stanford Encyclopedia of Philosophy* (Winter 2012 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/win2012/entries/two-dimensional-semantics/>>.
- Seager, W., 1999. *Theories of Consciousness*, Routledge.
- Searle, J., 1992. *The Rediscovery of Mind*, Cambridge, Mass: MIT Press.
- Searle, J., 1997. *The Mystery of Consciousness*, New York: The New York Review of Books.
- Shoemaker, S., 1982. "The Inverted Spectrum", *Journal of Philosophy*, 79: 357–81.
- Smullyan, R. M., 1981. "An epistemological nightmare", en Hofstadter, D. R. y D. C. Dennett ed., *The Mind's I: Fantasies and Reflections on Self and Soul*, Nueva York.
- Stoljar, D., 2007. "Consequences of Intentionalism", *Erkenntnis* (special issue), 66(1–2): 247–70.
- Stoljar, D., 2010. *Physicalism*, New York: Routledge.
- Strawson, P., 1959. *Individuals*, London: Methuen.
- Strawson, P., 1966. *The Bounds of Sense: an Essay on Kant's Critique of Pure Reason*, London and New York: Routledge.
- Stroud, B., 2000. *The Quest for Reality: Subjectivism and the Metaphysics of color*, New York: Oxford University Press.
- Tye, M., 1995. *Ten Problems of Consciousness*, Cambridge, Mass: MIT Press.
- Weiskrantz, B., 1986. *Blindsight. A Case Study and its Implications*, Oxford: Oxford University Press.
- Wertheimer, M., 1912. "Experimentelle Studien über das Sehen von Bewegung", *Zeitschrift für Psychologie*, 61, pp. 161-265.

Williams, B., 1978. *Descartes: The project of Pure Enquiry*, Harmondsworth: Penguin.

Williams, B., 2002. *Truth and Truthfulness*, Princeton, NJ y Oxford: Princeton University Press.

Williams, B., 2006. *Philosophy as a Humanistic Discipline*, Princeton y Oxford: Princeton University Press.